

SUMMARIES

Pred. No. is the number of results predicted by chance to have a

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	500	100.0	5199	6	BD205033	BD205033 Isolated
2	486	97.2	877	6	BD124639	BD124639 Primer to
3	486	97.2	877	6	BD126208	BD126208 Primer to
4	486	97.2	1299	9	BC039036	BC039036 Homo sapi
5	486	97.2	1676	6	BD127640	BD127640 Primer fo
6	486	97.2	1676	9	AK075290	AK075290 Homo sapi
7	486	97.2	2131	6	AX880753	AX880753 Sequence
8	486	97.2	2131	6	BD127168	BD127168 Primer to
9	486	97.2	2131	6	BD158570	BD158570 Primer fo
10	486	97.2	2131	9	AK072846	AK072846 Homo sapi
11	486	97.2	2131	9	AK074554	AK074554 Homo sapi
12	486	97.2	2199	9	BC058925	BC058925 Homo sapi
13	486	97.2	5195	6	BD205035	BD205035 Isolated
14	486	97.2	5381	9	AF114488	AF114488 Homo sapi
15	486	97.2	5458	6	BD205034	BD205034 Isolated
16	486	97.2	6439	9	AF114487	AF114487 Homo sapi
17	471.4	94.3	676	6	AX870076	AX870076 Sequence
18	471.4	94.3	676	6	BD150138	BD150138 Primer fo
19	428	85.6	2079	6	BD205036	BD205036 Isolated
20	399	79.8	5287	9	AF064243	AF064243 Homo sapi
21	399	79.8	7247	9	AF064244	AF064244 Homo sapi
22	280.2	56.0	3812	10	AF132672	AF132672 Rattus no
23	280.2	56.0	4025	10	AF132478	AF132478 Mus muscu
24	241.8	48.4	3723	10	AF132478	AF132478 Mus muscu
25	241.8	48.4	5145	10	AF132481	AF132481 Mus muscu
26	239.8	48.0	493	10	AF525079	AF525079 Mus muscu
27	216.6	43.3	4103	5	AF032118	AF032118 Xenopus l
28	206	41.2	3390	9	HSM809003	BX648852 Homo sapi
29	183	36.6	481	10	AF468654	AP000308 Homo sapi
30	163.2	32.6	37425	9	AP000308	AP000048 Homo sapi
31	163.2	32.6	57688	9	AP000048	AP0000191 Homo sapi
32	163.2	32.6	100000	9	AP000115	AP000115 Homo sapi
33	163.2	32.6	100000	9	AP000191	AP0001717 Homo sapi
34	161.6	32.3	340000	9	BS000182	BS000182 Pan trogl
35	161.6	32.3	163049	9	BS000181	BS000181 Pan trogl
36	161.6	32.3	199436	9	AC144363	AC144363 Papio anu
37	156.8	31.4	157138	2	BC020921	BC020921 Homo sapi
38	154.4	30.9	1749	9	BC038963	BC038963 Homo sapi
39	154.4	30.9	2085	9	AF182199	AF182199 Homo sapi
40	154.4	30.9	4557	9	AB033082	AB033082 Homo sapi
41	154.4	30.9	5938	9	AF182198	AF182198 Homo sapi
42	154.4	30.9	6091	9	BD167848	BD167848 Method fo
43	152.8	30.6	5828	6	AF248540	AF248540 Homo sapi
44	152.8	30.6	5828	6	BD264620	BD264620 Compositi
45	137.8	27.6	531	6	BD264620	

RESULT 1	LOCUS	DEFINITION
BD205033	BD205033	Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
	5199 bp	DNA linear PAT 17-JUL-2000

SOURCE ORGANISM	REFERENCE
Homo sapiens (human)	Korenbeld, J.R. and Chen, X.N. Isolated SH3 gene relating to myeloproliferative disorders and
Homo sapiens	1 (bases 1 to 5199)
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini, Homnidae; Homo.	

JOURNAL Patient: JP 2002511267-A/1.16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL
COMMENT OS Homo sapiens (human)
PN JP 2002511267-A/1
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JUIE R KORENBERG, XIAO NING CHEN
PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1..5199
FT Location/Qualifiers
1..5199
/organism='Homo sapiens (human)'.
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 100.0%; Score 500; DB 6; Length 5199;
Best Local Similarity 100.0%; Pred. No. 3.9e-120;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAAGATTCCGGGTACGGCGCTCGCGAGAGAAATCCCGAGCGGCTCCGGGACGGA 60
DB 1 CAAAAGATTCCGGGTACGGCGCTCGCGAGAGAAATCCCGAGCGGCTCCGGGACGGA 60
QY 61 CAGAGAGCGGCGGCGGATGCTGTGCGGCGCTCGGCTCTCTGCTCCCGACGGCGCG 120
DB 61 CAGAGAGCGGCGGCGGATGCTGTGCGGCGCTCGGCTCTCTGCTCCCGACGGCGCG 120
QY 121 GAGAGGCGGCGGATGCTGTGCGGCGCGGCGGAGCGCGCGGAGATGAGCGCTC 180
DB 121 GAGAGGCGGCGGATGCTGTGCGGCGCGGCGGAGCGCGCGGAGATGAGCGCTC 180
QY 181 GATTACCAAGTAAAGTAAAGAACCAATGCTCAGTTTCCAAACCTTTTGTGTCAGC 240
DB 181 GATTACCAAGTAAAGTAAAGAACCAATGCTCAGTTTCCAAACCTTTTGTGTCAGC 240
QY 241 CTGGATATCTGGGCGCTAACTGTAGAGAGAAAGCGGAGATGATGACAGTTCCATAGT 300
DB 241 CTGGATATCTGGGCGCTAACTGTAGAGAGAAAGCGGAGATGATGACAGTTCCATAGT 300
QY 301 TTAAGCCCAATATCTGATTCATTACTGGTGATCAAGCTAGAAACTTTTTTCAATCT 360
DB 301 TTAAGCCCAATATCTGATTCATTACTGGTGATCAAGCTAGAAACTTTTTTCAATCT 360
QY 361 GGGTTACCTCAACCTGTTTTTSCACAGATATGGGCACTAGCTGACATGAATATGATGA 420
DB 361 GGGTTACCTCAACCTGTTTTTSCACAGATATGGGCACTAGCTGACATGAATATGATGA 420
QY 421 AGAATGATCAATGAGATTTTCCATAGCTATGAATCTTATCAACTGAAGCTACAAAGA 480
DB 421 AGAATGATCAATGAGATTTTCCATAGCTATGAATCTTATCAACTGAAGCTACAAAGA 480
QY 481 TATCAGTACCTCTGCACT 500
DB 481 TATCAGTACCTCTGCACT 500

RESULT 2
LOCUS BD124639 877 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD124639
VERSION BD124639.1 GI:23219584

KEYWORDS JP 2002017375-A/70.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 877)
Oca,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.
Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patient: JP 2002017375-A 70 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/70
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OCA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA

FEATURES
source Location/Qualifiers
1..877
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 97.2%; Score 486; DB 6; Length 877;
Best Local Similarity 100.0%; Pred. No. 1.4e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCGAGAGAAATCCCGAGCGGCTCCGGGACGACAGAGAGCGGGCG 74
DB 12 GTACGGCGGCTCGCGAGAGAAATCCCGAGCGGCTCCGGGACGACAGAGAGCGGGCG 71
QY 75 GGGATGTGTCGGGCGGCTGCGGCTCTGCTCCCTCCAGCGCGGCTGAGCGGCACTGA 134
DB 72 GGGATGTGTCGGGCGGCTGCGGCTCTGCTCCCTCCAGCGCGGCTGAGCGGCACTGA 131
QY 135 TTTGTCCTGGGCGGCGAGCGCGGACCGCCGGAAGTGAAGGCTGATAGCAAGTAA 194
DB 132 TTTGTCCTGGGCGGCGAGCGCGGACCGCCGGAAGTGAAGGCTGATAGCAAGTAA 191
QY 195 AAGTACAGAACCATGCTCAGTTTCCAAACCTTTTGGTGGAGCTGATATCTGGGC 254
DB 192 AAGTACAGAACCATGCTCAGTTTCCAAACCTTTTGGTGGAGCTGATATCTGGGC 251
QY 255 CATACTGTAGAGAGAAAGCGAGCATGACAGCTTCATAGTTTAAAGCAATATC 314
DB 252 CATACTGTAGAGAGAAAGCGAGCATGACAGCTTCATAGTTTAAAGCAATATC 311
QY 315 TGGATTCATTCTGTTGATCAAGCTAGAAACTTTTTTGAATCTGGGTTACTCAAC 374
DB 312 TGGATTCATTCTGTTGATCAAGCTAGAAACTTTTTTGAATCTGGGTTACTCAAC 371
QY 375 TGTTTAGACAGATATGGGCACTAGCTGACATGAATATGAGAGATGATCAAGT 434
DB 372 TGTTTAGACAGATATGGGCACTAGCTGACATGAATATGAGAGATGATCAAGT 431
QY 435 GGAATTTCCATAGTATGAATCTTATCAACTGAAGCTACAGAGATATGACTACCTC 494
DB 432 GGAATTTCCATAGTATGAATCTTATCAACTGAAGCTACAGAGATATGACTACCTC 491
QY 495 TGCACCT 500

Db 492 TGCACT 497

|||||

RESULT 3

LOCUS BD126208 877 bp DNA linear PAT 18-SEP-2002

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD126208.1 GI:23221153

VERSION JP 2002017375-A/1639.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Makamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002017375-A 1639 22-JAN-2002;

COMMENT HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002017375-A/1639

PD 22-JAN-2002

PF 07-JUL-2000 JP 2000253172

PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO ISHII,

PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI SHINICHI KOJIMA,

PI TETSUJI OTSUKI, HISASHI KOGA

PC

C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

PC Primer for synthesizing full-length cDNA and use thereof FH Key

Location/Qualifiers

FT source 1..877 /organism="Homo sapiens (human)".

1..877 /mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 97.2%; Score 486; DB 6; Length 877;

Best Local Similarity 100.0%; Pred. No. 1.4e-116;

Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

15 GTACGGCGGCTCGCGAGAGATCCGAGCGGCTCCGGAGCGAGACGAGAGCGGCG 74

12 GTACGGCGGCTCGCGAGAGATCCGAGCGGCTCCGGAGCGAGACGAGAGCGGCG 71

75 GGGATGGGTGGCGGGCTCGCGGCTCTCGCTCCCGAGCGGCGCGAGAGCGGCGACTGA 134

72 GGGATGGGTGGCGGGCTCGCGGCTCTCGCTCCCGAGCGGCGCGAGAGCGGCGACTGA 131

135 TTTGTCCCTGGGCGGCGAGCGCGAGCCCGCGAGATGAGCGCTGCAATTAGCAAGGTTA 194

132 TTTGTCCCTGGGCGGCGAGCGCGAGCCCGCGAGATGAGCGCTGCAATTAGCAAGGTTA 191

195 AAGTAAACGAACCATGCTCAGTTTCCAAACCTTTTGGTGGAGCGCTGAGATATTTGGGC 254

192 AAGTAAACGAACCATGCTCAGTTTCCAAACCTTTTGGTGGAGCGCTGAGATATTTGGGC 251

255 CATAACTGTAGAGAAAGAGGAGCATGATGAGAGTTCATAGTTTAAAGCCATATTC 314

252 CATAACTGTAGAGAAAGAGGAGCATGATGAGAGTTCATAGTTTAAAGCCATATTC 311

315 TGGATTCTATTCTGTGTGATCAAGCTTAAAGCTTTTTCATCTGGGTTAAGCTCAAC 374

312 TGGATTCTATTCTGTGTGATCAAGCTTAAAGCTTTTTCATCTGGGTTAAGCTCAAC 371

QY 375 TGTTTAGCAAGATATGGGCACTAGCTGACATGATATATGAGAGATGATCAAGT 434

Db 372 TGTTTAGCAAGATATGGGCACTAGCTGACATGATATATGAGAGATGATCAAGT 431

QY 435 GAGATTTCCATAGCTATGAACTTATCAACTGAGCTAGACAGATATCACTACCCCTC 494

Db 432 GAGATTTCCATAGCTATGAACTTATCAACTGAGCTAGACAGATATCACTACCCCTC 491

QY 495 TGCACT 500

|||||

Db 492 TGCACT 497

RESULT 4

LOCUS BC039036 1299 bp mRNA linear PRI 24-DEC-2002

DEFINITION Homo sapiens, similar to intersectin 1 (SH3 domain protein), clone IMAGE:5784342, mRNA, partial cds.

ACCESSION BC039036

VERSION BC039036.1 GI:24659505

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (01-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA library Preparation: Life Technologies, Inc.

cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILNI)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nih.nih.gov

Akhter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, D., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, U.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: <http://image.llnl.gov>

Series: IRAP Plate: 82 Row: 1 Column: 2

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796.

Location/Qualifiers

1..1299

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5784342"

/tissue_type="uterus", leiomyosarcoma"

/clone_lib="NIH MGC_71"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

264..>1299

/codon_start=1

/product="Similar to intersectin 1 (SH3 domain protein)"

/protein_id="AAH39036.1"

/db_xref="GI:27371135"

CDS

/translation="AAOPPTPPGSLDIWALTVERBRKXDOOHSLKPISGFTGDQA
RMFPQSGPOPVLQIOMALAMNNDGDFEFSTAMLTIKIOGYOLPALPBPVAM
KOOPVIASSAPAPKAGIAMPPLUTVAVAPVMSITPVYKMSPTLUSVPTAAVPLAN
GAPVYQPLPAPAHAPATLTKSSFSRSGSOLNTKLPKAGSFVAVPVAEMAVP
OSSRLKRYQLPNSHDKTMSGHLTGAPATILMOSLPKQOLASIMLSIDIDGKTLA
EEFIIAMHILIDVAMSCQPLPVPVLPPEYIPPSFRVRSVSGISVISTSVDRLEPVP
LDEDOQKXKKK"

ORIGIN

Query Match 97.2%; Score 486; DB 9; Length 1299;
Best Local Similarity 100.0%; Pred. No. 1.5e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	15	GTACGGCGGCTGCGGAGAAAGATCCCGAGCGGGCTCCGGGAGCGACAGAGAGCGGGCG	74
DB	71	GTACGGCGGCTGCGGAGAAAGATCCCGAGCGGGCTCCGGGAGCGACAGAGAGCGGGCG	130
QY	75	GGGATGATGTGCGGGGCTGCGGCTCCGCTCCGCTCCAGCGGCGGTGAGCGGCACTGA	134
DB	131	GGGATGATGTGCGGGGCTGCGGCTCCGCTCCGCTCCAGCGGCGGTGAGCGGCACTGA	190
QY	135	TTTGTCCCTGGGGCGGCGAGCGGTGACCCGCGGAGATGAGCGGTGATTAGCAAGTAA	194
DB	191	TTTGTCCCTGGGGCGGCGAGCGGTGACCCGCGGAGATGAGCGGTGATTAGCAAGTAA	250
QY	195	AGGTACAGAACCATGCTCACTTTTCCACACCTTTTGGTGGCAGCGCTGGATATCTGGGG	254
DB	251	AGGTACAGAACCATGCTCACTTTTCCACACCTTTTGGTGGCAGCGCTGGATATCTGGGG	310
QY	255	CATTAAGTGAAGAAAGAGCGGAGAGATGATGAGAGTTCATGTTTAAAGCAATATC	314
DB	311	CATTAAGTGAAGAAAGAGCGGAGAGATGATGAGAGTTCATGTTTAAAGCAATATC	370
QY	315	TGATTCATTACCTGGTGATCAAGCTAGAAAATTTTTTTCATCTGGGTACTCAAC	374
DB	371	TGATTCATTACCTGGTGATCAAGCTAGAAAATTTTTTTCATCTGGGTACTCAAC	430
QY	375	TGTTTGAAGCAGATATGAGCACTAGCTGACATGAATTAATGAGAAAGATGATCAAGT	434
DB	431	TGTTTGAAGCAGATATGAGCACTAGCTGACATGAATTAATGAGAAAGATGATCAAGT	490
QY	435	GGAGTTTCCATATGCTATGAAACTTATCAAACTGAAGCTACAGAGATATGACTACCTC	494
DB	491	GGAGTTTCCATATGCTATGAAACTTATCAAACTGAAGCTACAGAGATATGACTACCTC	550
QY	495	TGCACT 500	
DB	551	TGCACT 556	

RESULT 5
BD127640 1676 bp DNA linear PAT 18-SEP-2002

LOCUS BD127640
DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD127640
VERSION BD127640.1 GI:23222585
KEYWORDS JP 2002017375-A/3071.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1676)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Makamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.

TITLE Journal
COMMENT Helix RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/3071
PD 22-JAN-2002 JP 2002253172
PF 07-JUL-2000 JP 2002253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO

PI ISHII,
PI YURI KAWAI, AI MAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUO OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12O1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
(264). (1676).
FT CDS Location/Qualifiers
1..1676
source /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 97.2%; Score 486; DB 6; Length 1676;
Best Local Similarity 100.0%; Pred. No. 1.5e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	15	GTACGGCGGCTGCGGAGAAAGATCCCGAGCGGGCTCCGGGAGCGACAGAGAGCGGGCG	74
DB	71	GTACGGCGGCTGCGGAGAAAGATCCCGAGCGGGCTCCGGGAGCGACAGAGAGCGGGCG	130
QY	75	GGGATGATGTGCGGGGCTGCGGCTCCGCTCCGCTCCAGCGGCGGTGAGCGGCACTGA	134
DB	131	GGGATGATGTGCGGGGCTGCGGCTCCGCTCCGCTCCAGCGGCGGTGAGCGGCACTGA	190
QY	135	TTTGTCCCTGGGGCGGCGAGCGGTGACCCGCGGAGATGAGCGGTGATTAGCAAGTAA	194
DB	191	TTTGTCCCTGGGGCGGCGAGCGGTGACCCGCGGAGATGAGCGGTGATTAGCAAGTAA	250
QY	195	AGGTACAGAACCATGCTCACTTTTCCACACCTTTTGGTGGCAGCGCTGGATATCTGGGG	254
DB	251	AGGTACAGAACCATGCTCACTTTTCCACACCTTTTGGTGGCAGCGCTGGATATCTGGGG	310
QY	255	CATTAAGTGAAGAAAGAGCGGAGAGATGATGAGAGTTCATGTTTAAAGCAATATC	314
DB	311	CATTAAGTGAAGAAAGAGCGGAGAGATGATGAGAGTTCATGTTTAAAGCAATATC	370
QY	315	TGATTCATTACCTGGTGATCAAGCTAGAAAATTTTTTTCATCTGGGTACTCAAC	374
DB	371	TGATTCATTACCTGGTGATCAAGCTAGAAAATTTTTTTCATCTGGGTACTCAAC	430
QY	375	TGTTTGAAGCAGATATGAGCACTAGCTGACATGAATTAATGAGAAAGATGATCAAGT	434
DB	431	TGTTTGAAGCAGATATGAGCACTAGCTGACATGAATTAATGAGAAAGATGATCAAGT	490
QY	435	GGAGTTTCCATATGCTATGAAACTTATCAAACTGAAGCTACAGAGATATGACTACCTC	494
DB	491	GGAGTTTCCATATGCTATGAAACTTATCAAACTGAAGCTACAGAGATATGACTACCTC	550
QY	495	TGCACT 500	
DB	551	TGCACT 556	

RESULT 6
AK075290 1676 bp mRNA linear PRI 03-SEP-2002

LOCUS AK075290
DEFINITION Homo sapiens cDNA FL90809 f1s, clone Y79AA1000778, highly similar
to Homo sapiens interseccin long form mRNA.

ACCESSION AK075290
VERSION AK075290.1 GI:22761283
KEYWORDS Oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahari, K., Masuno, Y., Ono, T., Okano, Y., Yoshikawa, Y.,
Aocuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1676)
Isogai, T. and Otsuki, T.
Direct Submision
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@helix.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo; Laboratory of
Genome Structure, Human Genome Center; cDNA 5'-3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1..1676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Y79AAl000778"
/cell_line="Y79"
/cell_type="retinoblastoma"
/clone_lib="Y79AAl"
/note="Cloning vector: pMEISFL3"

ORIGIN

Query Match 97.2%; Score 486; DB 9; Length 1676;
Best Local Similarity 100.0%; Pred. No. 1.5e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

15 GTACGGCGGCTCGCGAGGAAGATCCCGAGCGGCTCCGGGACGACAGAGAGGGGGG 74
71 GTACGGCGGCTCGCGAGGAAGATCCCGAGCGGCTCCGGGACGACAGAGAGGGGGG 130
75 GGGATGATGATGCGGGGCTGCGGCTCTGCTCCCGAGCGGCGGTGAGCGGCACTGA 134
131 GGGATGATGATGCGGGGCTGCGGCTCTGCTCCCGAGCGGCGGTGAGCGGCACTGA 190
135 TTTGTCCTGCGGGCGGACCGCGCCCGGAGATGAGCGGTGATTAGCAAGTAA 194
191 TTTGTCCTGCGGGCGGACCGCGCCCGGAGATGAGCGGTGATTAGCAAGTAA 250
195 AAGTAAACGAAACCATGCTCACTTTTGGTGGCAGCTTGGATATCTGGGC 254
251 AAGTAAACGAAACCATGCTCACTTTTGGTGGCAGCTTGGATATCTGGGC 310
255 CATTAAGTAAAGGAAGAGGAGCATGATGAGCATGCTTCAATGTTTAAAGCAATATC 314
311 CATTAAGTAAAGGAAGAGGAGCATGATGAGCATGCTTCAATGTTTAAAGCAATATC 370
315 TGGATTCATTACTGCGATCAAGTAAAGCTTTTTCATCTGGGTAACTCAACC 374
371 TGGATTCATTACTGCGATCAAGTAAAGCTTTTTCATCTGGGTAACTCAACC 430
375 TGTTTTGAAGACAGATATGGGCACTAGCTACATGATTAATGATGAAAGATGATTAAGT 434
431 TGTTTTGAAGACAGATATGGGCACTAGCTACATGATTAATGATGAAAGATGATTAAGT 490
435 GGAGTTTTCATAGCTATGAACCTTATCAACTGAAGTAAAGGATATCAGTACCTTC 494
491 GGAGTTTTCATAGCTATGAACCTTATCAACTGAAGTAAAGGATATCAGTACCTTC 550
495 TGCACCT 500
551 TGCACCT 556

RESULT 7
AX880753
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX880753 2131 bp DNA linear PAT 17-DEC-2003
Sequence 15658 from Patent EP1074617.
AX880753
AX880753.1 GI:40035489
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primers for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 15658 07-FEB-2001;
Research Association for Biotechnology (URP)

FEATURES

source
Location/Qualifiers
1..2131
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="unnamed protein product"
/codon_start=1
347..>2131
/protein_id="CAE92347.1"
/db_xref="GI:40035490"
/translation="MAQPTPEFGSLDIWATTEVERAKHQFSLKPISGITGDQA
RNFFQSGLPQVLAQIWMALADNMNDQVDFEISIAKLILKIQGYQPSALPPVM
KOQVLAISAAFAVAPPLANGAPVLPPLAFAPATLPASSSEFSGSSQANTK
LOKXQSPVAVPVAWEAVPQSSRLXKROLFNSHDKMSGLTNGPQRTIIQSSSLP
QAQASTWNLDIDDDKLTAEFTLNLHLDVMSGQPLPVLPPEYITPSPFRVRS
GSGISVISTVDQLEPEPVLDEQDQLEKLPVTFEDKRNPERNGQLEKRRQA
LLEORKEORLEAOLERABOERKERERORERKRLQLEKQLEKRELEQLEERBK
EIERREAKRELEROROLEWERNRORLEINQNRKQEDIVLAKAKYLEPELEALND
KHOLEGLQDIPRGLTQROEISTNSRELRLAIEITHLQOOLQESQOMGLIPEX
QINDOLKQVOQONSILHRSLVTLKRALAKELARQLMDQDEVETKRSKLQEDIF
NNQLKELREHNRKQLOKQKMEARLKO"

CDS

ORIGIN

Query Match 97.2%; Score 486; DB 6; Length 2131;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

15 GTACGGCGGCTCGCGAGGAAGATCCCGAGCGGCTCCGGGACGACAGAGAGGGGGG 74
154 GTACGGCGGCTCGCGAGGAAGATCCCGAGCGGCTCCGGGACGACAGAGAGGGGGG 213
75 GGGATGATGATGCGGGGCTGCGGCTCTGCTCCCGAGCGGCGGTGAGCGGCACTGA 134
214 GGGATGATGATGCGGGGCTGCGGCTCTGCTCCCGAGCGGCGGTGAGCGGCACTGA 273
218 TTTGTCCTGCGGGCGGACCGCGCCCGGAGATGAGCGGTGATTAGCAAGTAA 333
274 TTTGTCCTGCGGGCGGACCGCGCCCGGAGATGAGCGGTGATTAGCAAGTAA 393
135 TTTGTCCTGCGGGCGGACCGCGCCCGGAGATGAGCGGTGATTAGCAAGTAA 194
191 TTTGTCCTGCGGGCGGACCGCGCCCGGAGATGAGCGGTGATTAGCAAGTAA 254
195 AAGTAAACGAAACCATGCTCACTTTTGGTGGCAGCTTGGATATCTGGGC 254
251 AAGTAAACGAAACCATGCTCACTTTTGGTGGCAGCTTGGATATCTGGGC 314
255 CATTAAGTAAAGGAAGAGGAGCATGATGAGCATGCTTCAATGTTTAAAGCAATATC 314
311 CATTAAGTAAAGGAAGAGGAGCATGATGAGCATGCTTCAATGTTTAAAGCAATATC 374
315 TGGATTCATTACTGCGATCAAGTAAAGCTTTTTCATCTGGGTAACTCAACC 374
371 TGGATTCATTACTGCGATCAAGTAAAGCTTTTTCATCTGGGTAACTCAACC 434
375 TGTTTTGAAGACAGATATGGGCACTAGCTACATGATTAATGATGAAAGATGATTAAGT 434
431 TGTTTTGAAGACAGATATGGGCACTAGCTACATGATTAATGATGAAAGATGATTAAGT 494
435 GGAGTTTTCATAGCTATGAACCTTATCAACTGAAGTAAAGGATATCAGTACCTTC 550
491 GGAGTTTTCATAGCTATGAACCTTATCAACTGAAGTAAAGGATATCAGTACCTTC 573
495 TGCACCT 500
551 TGCACCT 556

```

Db      574 GGAATTTCCATACCTATGAAACTTATCAAACTGAGTACAGAGATATGAGTACCTC 633
QY      495 TGCACT 500
Db      634 TGCACT 639

RESULT 8
BD127168
LOCUS
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127168
VERSION BD127168.1 GI:23222113
KEYWORDS JP 2002017375-A/2599.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          1 (bases 1 to 2131)
          Oca,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
          Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
          Koga,H.
          Primer for synthesizing full-length cDNA and use thereof
          Patent: JP 2002017375-A 2599 22-JAN-2002;
          HELIX RESEARCH INSTITUTE
          OS Homo sapiens (human)
          PN JP 2002017375-A/2599
          PD 22-JAN-2002
          PF 07-JUL-2000 JP 2000253172
          PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
          PI ISHII,
          PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
          SHINICHI KOJIMA,
          PI TETSUO OTSUKI,HISASHI KOGA
          PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
          10,
          PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
          Primer for synthesizing full-length cDNA and use thereof FH key
          Location/Qualifiers
          FT CDS
          source
          1..2131
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"

ORIGIN
Query Match 97.2%; Score 486; DB 6; Length 2131;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      15 GTACGGCGGCTCGCGAGAAATCCCGAGCGGCTCCGGAGCGAGACAGAGCGGGCG 74
Db      12 GTACGGCGGCTCGCGAGAAATCCCGAGCGGCTCCGGAGCGAGACAGAGCGGGCG 71
QY      75 GGGATGATGTCGGGGGCTGCGCTCTGCTCCCTCCACGGCGGTGAGGGCGCACTGA 134
Db      72 GGGATGATGTCGGGGGCTGCGCTCTGCTCCCTCCACGGCGGTGAGGGCGCACTGA 131
QY      135 TTTGTCCTGCGGGCGGCGAGCGCGACCGCGCGAGATGAGCGCTGATTAAGCAAGTAA 194
Db      132 TTTGTCCTGCGGGCGGCGAGCGCGACCGCGCGAGATGAGCGCTGATTAAGCAAGTAA 191
QY      195 AAGTAAACAACCATGCTCACTTTTCAACACTTTTGGTGGCAAGCTGATATCTGGGC 254
Db      192 AAGTAAACAACCATGCTCACTTTTCAACACTTTTGGTGGCAAGCTGATATCTGGGC 251
QY      255 CATAACTGTAGAGGAAGAAGCGCAAGCATGATGAGAGTTTCAATGTTTAAAGCAATATC 314
Db      252 CATAACTGTAGAGGAAGAAGCGCAAGCATGATGAGAGTTTCAATGTTTAAAGCAATATC 311

```

```

QY      315 TGATTCATTACTGTGATGATCAAGTAGAACTTTTTCATCTGGGTACTCAACC 374
Db      312 TGATTCATTACTGTGATGATCAAGTAGAACTTTTTCATCTGGGTACTCAACC 371
QY      375 TGTTTAGCAGATATGAGGCACTAGCTGACATGATATGATGAGAAATGATCAAGT 434
Db      372 TGTTTAGCAGATATGAGGCACTAGCTGACATGATATGATGAGAAATGATCAAGT 431
QY      435 GGAATTTCCATACCTATGAAACTTATCAAACTGAGCTACAGAGATATGACTACCTC 494
Db      432 GGAATTTCCATACCTATGAAACTTATCAAACTGAGCTACAGAGATATGACTACCTC 491
QY      495 TGCACT 500
Db      492 TGCACT 497

RESULT 9
BD158570
LOCUS
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD158570
VERSION BD158570.1 GI:27864328
KEYWORDS JP 2002191363-A/13413.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          1 (bases 1 to 2131)
          Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
          Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
          Primer for synthesizing full-length cDNA and use thereof
          Patent: JP 2002191363-A 13413 09-JUL-2002;
          HELIX RESEARCH INSTITUTE
          OS Homo sapiens (human)
          PN JP 2002191363-A/13413
          PD 09-JUL-2002
          PF 28-JUL-2000 JP 2000280990
          PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
          PI SATO,
          PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
          PI KEIICHI NAGAI,TETSUO OTSUKI
          PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
          10,
          PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
          Primer for synthesizing full-length cDNA and use thereof FH key
          Location/Qualifiers
          FT CDS
          source
          1..2131
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"

ORIGIN
Query Match 97.2%; Score 486; DB 6; Length 2131;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      15 GTACGGCGGCTCGCGAGAAATCCCGAGCGGCTCCGGAGCGAGACAGAGCGGGCG 74
Db      154 GTACGGCGGCTCGCGAGAAATCCCGAGCGGCTCCGGAGCGAGACAGAGCGGGCG 213
QY      75 GGGATGATGTCGGGGGCTGCGCTCTGCTCCCTCCACGGCGGTGAGGGCGCACTGA 134
Db      214 GGGATGATGTCGGGGGCTGCGCTCTGCTCCCTCCACGGCGGTGAGGGCGCACTGA 273
QY      135 TTTGTCCTGCGGGCGGCGAGCGCGACCGCGCGAGATGAGCGCTGATTAAGCAAGTAA 194
Db      274 TTTGTCCTGCGGGCGGCGAGCGCGACCGCGCGAGATGAGCGCTGATTAAGCAAGTAA 333
QY      195 AAGTAAACAACCATGCTCACTTTTCAACACTTTTGGTGGCAAGCTGATATCTGGGC 254

```

Db 334 AAGTAACTGAGAACCTGGCTCAAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGC 393
 QY 255 CATAACTGTAGAGGAAAGAGCGAGCATGATCAGACAGTTTCCATAGTTTAAAGCCAAATATC 314
 Db 334 CATACCTGTAGAGGAAAGAGCGAGCATGATCAGACAGTTTCCATAGTTTAAAGCCAAATATC 453
 QY 315 TGGATTCTATTACTGGTGTATCAAGCTAGAAACCTTTTTCATCTGGGTTACCTCAACC 374
 Db 454 TGGATTCTATTACTGGTGTATCAAGCTAGAAACCTTTTTCATCTGGGTTACCTCAACC 513
 QY 375 TGGTTAGACAGATATGGGAGCTAGCTGACATGATATATATATGAGAGATGATCAAGT 434
 Db 514 TGGTTAGACAGATATGGGAGCTAGCTGACATGATATATATGAGAGATGATCAAGT 573
 QY 435 GGAGTTTTCATAGCTATGAAACTTATCAAACTGAAGCTACAGATATCACTACCTC 494
 Db 574 GGAGTTTTCATAGCTATGAAACTTATCAAACTGAAGCTACAGATATCACTACCTC 633
 QY 495 TGCACCT 500
 Db 634 TGCACCT 639
 RESULT 10
 AK027846 2131 bp mRNA linear PRI 01-AUG-2002
 LOCUS Homo sapiens cDNA FLJ14940 fis, clone PLACE1010942, highly similar
 DEFINITION to Homo sapiens intersectin long isoform (ITSN) mRNA.
 ACCESSION AK027846
 KEYWORDS oligo capping, fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1
 AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
 Nishikawa, T., Nagai, K., Sugano, S., Ishibashi, T., Fujimori, K.,
 Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Ishii, S., Kawai, Y.,
 Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K.,
 Masuo, Y., and Kanehori, K.
 COMMENT NEDO human cDNA sequencing project
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 2131)
 REFERENCE Isogai, T. and Otsuki, T.
 AUTHORS Direct Submission
 JOURNAL Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology; cDNA library construction;
 5'- & 3'-end one pass sequencing and clone selection; Helix
 Research Institute (supported by Japan Key Technology Center etc.)
 and Department of Virology, Institute of Medical Science,
 University of Tokyo.
 FEATURES
 source
 Location/Qualifiers
 1..2131
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1010942"
 /tissue_type="Placenta"
 /clone_lib="PLAC1"
 /note="cloning vector: pME186FL3"
 ORIGIN
 Query Match 97.2%; Score 486; DB 9; Length 2131;
 Best Local Similarity 100.0%; Pred. No. 1.6e-116;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 15 GTACGGCGGCTCGCAGGAAGATCCCGAGCGGGCTCCGGAGCGGACAGAGAGCGGGCG 74

Db 154 GTACGGCGGCTCGCAGGAAGATCCCGAGCGGGCTCCGGAGCGGACAGAGAGCGGGCG 213
 QY 75 GGATGTGTGTGCGGGGCTGCGGCTCTGCTGCTCCAGCGGCGGTGAGCGGCACTGA 134
 Db 214 GGATGTGTGTGCGGGGCTGCGGCTCTGCTGCTCCAGCGGCGGTGAGCGGCACTGA 273
 QY 135 TTTGTCCCTGGGGCGGCGGCGGAGCCCGCCCGGAGATGAGGCGTCATTTAGCAAGTGA 194
 Db 274 TTTGTCCCTGGGGCGGCGGCGGAGCCCGCCCGGAGATGAGGCGTCATTTAGCAAGTGA 333
 QY 195 AAGTAACTGAGAACCTGGCTCAAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGC 254
 Db 334 AAGTAACTGAGAACCTGGCTCAAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGC 393
 QY 255 CATAACTGTAGAGGAAAGAGCGAGCATGATCAGACAGTTTCCATAGTTTAAAGCCAAATATC 314
 Db 334 CATAACTGTAGAGGAAAGAGCGAGCATGATCAGACAGTTTCCATAGTTTAAAGCCAAATATC 453
 QY 315 TGGATTCTATTACTGGTGTATCAAGCTAGAAACCTTTTTCATCTGGGTTACCTCAACC 374
 Db 454 TGGATTCTATTACTGGTGTATCAAGCTAGAAACCTTTTTCATCTGGGTTACCTCAACC 513
 QY 375 TGGTTAGACAGATATGGGAGCTAGCTGACATGATATATATGAGAGATGATCAAGT 434
 Db 514 TGGTTAGACAGATATGGGAGCTAGCTGACATGATATATGAGAGATGATCAAGT 573
 QY 435 GGAGTTTTCATAGCTATGAAACTTATCAAACTGAAGCTACAGATATCACTACCTC 494
 Db 574 GGAGTTTTCATAGCTATGAAACTTATCAAACTGAAGCTACAGATATCACTACCTC 633
 QY 495 TGCACCT 500
 Db 634 TGCACCT 639
 RESULT 11
 AK074554 2131 bp mRNA linear PRI 03-SEP-2002
 LOCUS Homo sapiens cDNA FLJ90073 fis, clone HEMBA1004110, highly similar
 DEFINITION to Homo sapiens intersectin short form mRNA.
 ACCESSION AK074554
 KEYWORDS oligo capping, fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1
 AUTHORS Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
 Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
 Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
 Kojima, S., Nagahara, K., Masuo, Y., Ono, T., Okano, K., Yoshikawa, Y.,
 Aotake, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
 Nimomiya, K.
 COMMENT NEDO human cDNA sequencing project
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 2131)
 REFERENCE Isogai, T. and Otsuki, T.
 AUTHORS Direct Submission
 JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology; cDNA library construction;
 Institute of Medical Science, University of Tokyo, Laboratory of
 Genome Structure, Human Genome Center, cDNA 5'- & 3'-end one pass
 sequencing and clone selection; Helix Research Institute (supported
 by Japan Key Technology Center etc.).
 FEATURES
 source
 Location/Qualifiers
 1..2131
 /organism="Homo sapiens"

misc_feature
335..619
/gene="TTSN1"
/note="EH; Region: Eps15 homology domain"
/db_xref="CDD:smart0027"

misc_feature
935..1171
/gene="TTSN1"
/note="EH; Region: Eps15 homology domain"
/db_xref="CDD:smart0027"

ORIGIN

Query Match 97.2%; Score 486; DB 9; Length 2199;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCCGGACGAGACGAGAGCGGCGG 74
DB 103 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCCGGACGAGAGCGGCGGCGG 162
QY 75 GGATGGTGTGGGGGCTCGCGGCTCGGCTCCCGAGCGGCGCGTGAAGCGGCACTGA 134
DB 163 GGATGGTGTGGGGGCTCGCGGCTCGGCTCCCGAGCGGCGCGTGAAGCGGCACTGA 222
QY 135 TTGTCCCTGGGGGCGGCGGCGGCGGCGGCGGAGATGAGCGCTGATTAAGCAAGTAA 194
DB 223 TTGTCCCTGGGGGCGGCGGCGGCGGCGGCGGAGATGAGCGCTGATTAAGCAAGTAA 282
QY 195 AAGTACGAGAACCATGCTGCTCAATCTTTCGAGCTTTGGTGGAGCTTGATATTGGGC 254
DB 283 AAGTACGAGAACCATGCTGCTCAATCTTTCGAGCTTTGGTGGAGCTTGATATTGGGC 342
QY 255 CATACCTGAGAGAGAGAGAGAGATGATGACGATTTCAATTTAAAGCCAAATTC 314
DB 343 CATACCTGAGAGAGAGAGAGAGATGATGACGATTTCAATTTAAAGCCAAATTC 402
QY 315 TGGATTCTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTAAGCTCAAC 374
DB 403 TGGATTCTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTAAGCTCAAC 462
QY 375 TGTTTTACACAGATATGGGCACTAGCTGACATGATATATGATGAAAGATGATCAAGT 434
DB 463 TGTTTTACACAGATATGGGCACTAGCTGACATGATATATGATGAAAGATGATCAAGT 522
QY 435 GGAGTTTCCATAGCTATGAACCTATCAAACTGAAGCTACAGGATTCAGCTACCTC 494
DB 523 GGAGTTTCCATAGCTATGAACCTATCAAACTGAAGCTACAGGATTCAGCTACCTC 582
QY 495 TGCACCT 500
DB 583 TGCACCT 588

RESULT 13
BD205035 5195 bp DNA linear PAT 17-JUL-2003
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205035
VERSION BD205035.1 GI:33014805
KEYWORDS JP 2002511267-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5195)
AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 3 16-APR-2002;

COMMENT
OS CEDARS SINAI HEALTH SYSTEM ET AL
PN OS Homo sapiens (human)
PD JP 2002511267-A/3
PF 16-APR-2002
PR 16-APR-1999 JP 2000543610
PI 16-APR-1998 US 60/082007
PC JULIE R. KORENBERG, XING NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12N13/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
and utilization thereof.
FH Key Location/Qualifiers
FT source 1..5195
FT FT /organism="Homo sapiens (human)".
FEATURES
source
1..5195
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 97.2%; Score 486; DB 6; Length 5195;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCCGGACGAGACGAGAGCGGCGG 74
DB 46 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCCGGACGAGACGAGAGCGGCGG 105
QY 75 GGATGGTGTGGGGGCTCGCGGCTCGGCTCCCGAGCGGCGCGTGAAGCGGCACTGA 134
DB 106 GGATGGTGTGGGGGCTCGCGGCTCGGCTCCCGAGCGGCGCGTGAAGCGGCACTGA 165
QY 135 TTGTCCCTGGGGGCGGCGGCGGCGGCGGCGGAGATGAGCGCTGATTAAGCAAGTAA 194
DB 166 TTGTCCCTGGGGGCGGCGGCGGCGGCGGCGGAGATGAGCGCTGATTAAGCAAGTAA 225
QY 195 AAGTACGAGAACCATGCTGCTCAATCTTTCGAGCTTTGGTGGAGCTTGATATTGGGC 254
DB 226 AAGTACGAGAACCATGCTGCTCAATCTTTCGAGCTTTGGTGGAGCTTGATATTGGGC 285
QY 255 CATACCTGAGAGAGAGAGAGAGATGATGACGATTTCAATTTAAAGCCAAATTC 314
DB 286 CATACCTGAGAGAGAGAGAGAGATGATGACGATTTCAATTTAAAGCCAAATTC 345
QY 315 TGGATTCTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTAAGCTCAAC 374
DB 346 TGGATTCTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTAAGCTCAAC 405
QY 375 TGTTTTACACAGATATGGGCACTAGCTGACATGATATATGATGAAAGATGATCAAGT 434
DB 406 TGTTTTACACAGATATGGGCACTAGCTGACATGATATATGATGAAAGATGATCAAGT 465
QY 435 GGAGTTTCCATAGCTATGAACCTATCAAACTGAAGCTACAGGATTCAGCTACCTC 494
DB 466 GGAGTTTCCATAGCTATGAACCTATCAAACTGAAGCTACAGGATTCAGCTACCTC 525
QY 495 TGCACCT 500
DB 526 TGCACCT 531

RESULT 14
AF114488 5381 bp mRNA linear PRI 16-JUL-2002
LOCUS AF114488
DEFINITION Homo sapiens intersectin short isoform (TTSN) mRNA, complete cds.
ACCESSION AF114488
VERSION AF114488.1 GI:4808824
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM	Homo sapiens
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS	1 (bases 1 to 5381)
TITLE	Pucharcos, C., Fuentes, J., Casas, C., de la Luna, S., Alcantara, S., Arbones, M., Soriano, E., Estivill, X., and Pritchard, M. Alu-splice cloning of human Interecclin (ITSN), a putative multivalent binding protein expressed in proliferating and differentiating neurons and overexpressed in Down syndrome
JOURNAL	Eur. J. Hum. Genet. 7 (6), 704-712 (1999)
MEDLINE	99415290
PUBMED	10482960
REFERENCE	2 (bases 1 to 5381)
AUTHORS	Pucharcos, C., Fuentes, J., Pritchard, M., and Estivill, X.
TITLE	Direct Submission
JOURNAL	Submitted (16-DEC-1998) Medical and Molecular Genetics Center, Cancer Research Institute, L'Hospitalet de Llo., Avia. Castelldefels km. 2,7, Barcelona 08907, Spain
FEATURES	Location/Qualifiers
source	1..5381
gene	1..5381
CDS	269..3931
	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="21" /map="21q22.1-22.2" /gene="ITSN" /gene="ITSN" /codon_start=1 /product="Interecclin short isoform" /protein_id="AAD29953.1" /db_xref="GI:4808825" /translation="MAQPPFGSSLDIATTVVERAKHDQFSLKPSIGFTIGDGO RNFPSGLPQPVLAQIMALADMMNDGMDQVSIAMKILIKLOGVLPALPVPW KQGVATSSAPRGMGGLASMPRTNAPVPMGSIPIVGMSPITVSSIFTPAVPILAN GAPVITQPPAPAPPAATLPEKSSFSFSSPESQSLNTRKQKQSDVASVPVAAVAVH OSSLYKRYQLPNSHDKTMGHLTGPOATLIMQSLPAQASLWNLSDIDQDKTLTA EEFILAMHLIDVAMSGQPLPVLPPPEVYIPSEFVRVSGSGISVTSVQRLPEEPV LEEDQQLLEKLIPTYPEDKRENFENGRLLEKFORALBOGRKORLALERAEOE KRKEEROERKRLLELEKQERLEORERERREIEREAKELEROLWME RNRQELINLRNKEDEDIIVYLKAKKTLPEFLALNKHQLEKLDIQRLLTQGC EISTNSKRELRLSIFTHLQQLDSQMLERLIPKQILNDQLKQVQNSLRSLD TLKALAKELARQLHEDQDLDEVEKETRSKLQDITFNNQLKEIRLHNKQQLQKXS MEARLQKEQERKIIELEKQEAQRAQBGADOMLHVQEDHORPRKLHEEKV KRRESVKKQSEERKQEOADQKGLRFLPHQEPAPKAVQAPWSTAEGPLTISAENV KVVYRALYPESSRSHDEITLOPDIIVWKGEMVDESOTGPGWLGSLKGTGMFP NYAEKIPENEPVAPVYKVDSTSAAPAKLALRETPALPATVATSSSPPTPNMADPSST WPTSTNEKPEPTDMDAMAAPSLTVPNSGQLRQSAATPATATSSSPVLAGSEKVB GLQALATYPRAKKNDNLNFKNDVITVLEQDDMMFGEVQGGMPKSPVKLISG IRTKSDMSGSSSEPSALKRVASPAAPVVSGEFIAMTYESSEQDLTFVQGDVILI VTKKDDGMWTSITVDKAGVPSNVYRLKDSGSGTAGTGLGKKPELAQVATSYTAN GPEDLTAAPQGLILIRKKNPEGWMEGLQAGKRQIGMFPANVKLISPEKSTLPTFL EPEPTSLAAQCVITGMVDTAQNDDIELAFKQIINVLNKEDPDMMKGEVNGVGLF PSNVVKLTTMDPESQ"
ORIGIN	
Query Match	97.2%; Score 486; DB 9; Length 5381;
Best Local Similarity	100.0%; Prid. No. 1,8e-116;
Matches	486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	15 GTACGGGGGCTCCGAGAGAGAAATCCGAGAGGGGCTCCGGAGCAGACAGAGGCGGGCG 74
DB	76 GTACGGCGGGCTCCGAGAGAGAAATCCGAGAGGGGCTCCGGAGCAGACAGAGGCGGGCG 135
QY	75 GGGATGGTGTGTCGGGGGTGCGGCTCTGTCGCTCCCTCCAGCGAGCGCGTGAAGCGGCACATGA 134
DB	136 GGGATGGTGTGTCGGGGGTGCGGCTCTGTCGCTCCCTCCAGCGAGCGCGTGAAGCGGCACATGA 195
QY	135 TTGTGTCTCTGGGGGGGCGAGCGCGGACCCCGCCGAGATGAGGCTGTGATTAGCAAGTAA 194
DB	196 TTGTGTCTCTCTGGGGGGGCGAGCGCGGACCCCGCCGAGATGAGGCTGTGATTAGCAAGTAA 255

QY	195	AAGTAAACAGAACCATGGCTCGAGTTTCCAAACCTTTTGGTGGCAGCCTGATATCTGGGC	254
Db	256	AAGTAAACAGAACCATGGCTCGAGTTTCCAAACCTTTTGGTGGCAGCCTGATATCTGGGC	315
QY	255	CATAACTGTAGAGGAAAGAGGAGCAGTATGATCAGCAGTTCATGTTTAAAGCCAAATATC	314
Db	316	CATAACTGTAGAGGAAAGAGGAGCAGTATGATCAGCAGTTCATGTTTAAAGCCAAATATC	375
QY	315	TGAGTTCAATTACTGGTGTCAAGCTAGAACTTTTTCATCTGGGTTTACTCTAAC	374
Db	376	TGAGTTCAATTACTGGTGTCAAGCTAGAACTTTTTCATCTGGGTTTACTCTAAC	435
QY	375	TGTTTTCAGCAGATATGGGCACTAGCTGCATGATATATGATGAGGAATGATCAAGT	434
Db	436	TGTTTTCAGCAGATATGGGCACTAGCTGCATGATATATGATGAGGAATGATCAAGT	495
QY	435	GGAGTTTTCATAGCTATGAAACTTATCAACTGAGCTACCAAGATATACGTAACCTC	494
Db	496	GGAGTTTTCATAGCTATGAAACTTATCAACTGAGCTACCAAGATATACGTAACCTC	555
QY	495	TGCACT 500	
Db	556	TGCACT 561	
RESULT 15			
LOCUS	BD205034		
DEFINITION	Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.		
ACCESSION	BD205034		
VERSION	BD205034.1	GI:33014804	
KEYWORDS	JP 2002511267-A/2.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi, Mammalia, Eutheria, Primates, Catarrhini, Homnidae, Homo.		
AUTHORS	1 (bases 1 to 5458)		
TITLE	Korenberg, J.R. and Chen, X.N.		
JOURNAL	Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof		
COMMENT	Patent: JP 2002511267-A 2 16-APR-2002; CEDARS SINAI HEALTH SYSTEM ET AL		
	OS Homo sapiens (human)		
	PN JP 2002511267-A/2		
	PD 16-APR-2002		
	PF 16-APR-1999 JP 2000543610		
	PR 16-APR-1998 US 60/082007		
	PT JULIE R KORENBERG, XIAO NING CHEN		
	PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,		
	PC C1201/68,		
	PC G01N33/68//A61K48/00, C12N15/00, C12N5/00		
	CC Isolated SH3 gene relating to myeloproliferative disorders and		
	leukemia		
	CC and utilization thereof.		
	CC Key Location/Qualifiers		
	FT source 1..5458		
	FT location/Qualifiers /Organism='Homo sapiens (human)'. 1..5458		
	location/Qualifiers /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"		
FEATURES			
SOURCE			
ORIGIN			
Query Match	97.2%; Score 486; DB 6; Length 5458;		
Best Local Similarity	100.0%; Pred No. 1.8e-116;		
Matches	486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	15	GTAACGGCGGCTCCGAGGAAGATCCCGAGCGGCTCCGGACGGAAGAAGCGCGGCG	74
Db	74	GTAACGGCGGCTCCCGAGGAAGATCCCGAGCGGCTCCGGACGGAAGAAGCGCGGCG	133

```
QY 75 GGGATGCTGTGCGGGGCTGCGGCTCTGCGTCCCTCCAGCGCGCGGTGAGCGGCATCTGA 134
    |||||
Db 134 GGGATGCTGTGCGGGGCTGCGGCTCTGCGTCCCTCCAGCGCGCGGTGAGCGGCATCTGA 193
    |||||
QY 135 TTTGTCCCTGCGGCGCGGCGCGGCGCGGCGGATGAGGCGGTGATTTAGCAAGGTAA 194
    |||||
Db 194 TTTGTCCCTGCGGCGCGGCGCGGCGCGGCGGATGAGGCGGTGATTTAGCAAGGTAA 253
    |||||
QY 195 AAGTAACGAACCATGCGTCAAGTTCCAAACCTTTTGTGCGAGCCTGGAATATCTGCGC 254
    |||||
Db 254 AAGTAACGAACCATGCGTCAAGTTCCAAACCTTTTGTGCGAGCCTGGAATATCTGCGC 313
    |||||
QY 255 CATTAAGTGTAGAGGAAAGCGGAAGCATGATCAGCAGTTCCATAGTTTAAAGCCATATC 314
    |||||
Db 314 CATTAAGTGTAGAGGAAAGCGGAAGCATGATCAGCAGTTCCATAGTTTAAAGCCATATC 373
    |||||
QY 315 TGGATTCATTTACTGCTGATCAAGCTAGAAGCTTTTTCATCTGGGTTACCTCAACG 374
    |||||
Db 374 TGGATTCATTTACTGCTGATCAAGCTAGAAGCTTTTTCATCTGGGTTACCTCAACG 433
    |||||
QY 375 TGTTTTACACAGATATGCGGCACTAGTGAATGAATATGATGAGAGATGATCAAGT 434
    |||||
Db 434 TGTTTTACACAGATATGCGGCACTAGTGAATGAATATGATGAGAGATGATCAAGT 493
    |||||
QY 435 GGAAGTTTCCATAGCTATGAAGCTTATCAAACTGAAGCTACAGAGATATCAGTACCTC 494
    |||||
Db 494 GGAAGTTTCCATAGCTATGAAGCTTATCAAACTGAAGCTACAGAGATATCAGTACCTC 553
    |||||
QY 495 TGCAC 500
    |||||
Db 554 TGCAC 559
```

Search completed: July 1, 2004, 12:16:42
Job time : 2138 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2004, 10:19:06 ; Search time 2656 Seconds

(without alignments)
5621.645 Million cell updates/sec

Title: US-09-720-934-1_COPY_1_500

Perfect score: 500
Sequence: 1 caaagaatccgggtacgg.....taccagctaccctctgact 500

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_esthum:*
4: em_esthum:*
5: em_esthum:*
6: em_esthum:*
7: em_esthum:*
8: em_esthum:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_esthum:*
16: em_esthum:*
17: em_gss_hum:*
18: em_gss_hum:*
19: em_gss_hum:*
20: em_gss_hum:*
21: em_gss_hum:*
22: em_gss_hum:*
23: em_gss_hum:*
24: em_gss_hum:*
25: em_gss_hum:*
26: em_gss_hum:*
27: em_gss_hum:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	486	97.2	896	13	BQ941336 AGENCOURT
2	486	97.2	1089	12	BQ050397 AGENCOURT
3	484.4	96.9	979	13	BQ941411 AGENCOURT
4	474	94.8	750	12	BG118422 602348040

5	474	94.8	2126	11	BC020269 Homo sapi
6	460	92.0	643	13	BX470886 DKFZP686E
7	459.8	92.0	811	12	BG829540 602763842
8	454	90.8	544	14	AL711737 DKFZP686K
9	437.4	87.5	605	14	CB265595 12822043
10	424	84.8	2079	11	BC013578 Homo sapi
11	413	82.6	734	14	CB269754 1008661 H
12	397.6	79.5	405	9	AL750406
13	387.6	77.5	547	10	BF194430 246586 MA
14	367.4	73.5	698	14	CB519152
15	367.4	73.5	5385	11	BC062938 Mus muscu
16	366.2	73.2	723	10	BE786696 601475048
17	363.8	72.8	624	13	BY713771 BY713771
18	329.6	65.9	662	14	CF536420 UT-M-GTO-
19	323.2	64.6	602	14	CA528645 8082-31 M
20	298.2	59.6	600	14	CA529211 8101-57 M
21	297	59.4	411	9	AL549192 UT-R-C3-E
22	293.2	58.6	475	10	BE199657 ug70a11.x
23	293.2	58.6	488	9	AL481705 v330f01.x
24	292.2	58.4	445	9	AL852079 UT-M-BHO-
25	292.2	58.4	501	9	AL852070 UT-M-BHO-
26	292.2	58.4	510	10	BE952536 UT-M-CRO-
27	291.8	58.4	394	12	BM484041 537738 MA
28	289	57.8	359	9	AL853008 UT-M-BHO-
29	283.2	55.6	452	10	BB840012 603860857
30	276.8	55.4	523	13	BU381021 603860857
31	271.4	54.3	533	29	CG579249 OSM217389
32	270	54.0	721	13	BK470913 DKFZP686G
33	267.4	53.5	728	13	BY731854 BY731854
34	266.4	53.3	766	14	CF743758 UT-M-GIO-
35	254	50.8	477	14	CB269094 1008001 H
36	249.8	50.0	907	13	BU128229 603114019
37	246.8	49.4	873	14	CF728634 UT-M-HDO-
38	245.2	49.0	298	10	BE103413 UT-R-BXO-
39	244.6	48.9	299	10	AM890668 QV1-NT004
40	232.8	46.6	643	13	BU406611 604136611
41	218.2	43.6	589	9	AL792490 AL792490
42	215.6	43.3	534	12	BL660235 de42d08.Y
43	215.6	43.1	550	12	BL623100 BL623100
44	215	43.0	504	9	AL964682 AL964682
45	196.6	39.3	266	10	BE954163 UT-M-CDI-

ALIGNMENTS

RESULT 1
LOCUS BQ941336
DEFINITION AGENCOURT_8741326 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6420600
ACCESSION BQ941336
VERSION BQ941336.1 GI:22356814
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://imgc.ncbi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DP/GenDat
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at:
http://image.llnl.gov
Plate: LCM2594 row: k column: 01
High quality sequence stop: 763.

FEATURES		Location/Qualifiers
source	1..896	/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="IMAGE:5420600"
		/tissue_type="large cell carcinoma"
		/lab_host="DH10B (phage-resistant)"
		/clone_11b="NIH MGC 18"
		/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."
ORIGIN		
Query Match	97.2%; Score 486; DB 13; Length 896;	
Best Local Similarity	100.0%; Pred. No. 7.3e-116;	
Matches 486;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	15	GTACGCGCGCTTCGCCAGAGAAATATCCGAGCGGCTCCGCGACGACAGAGAGTGGGCG 74
DB	17	GTACGCGCGCTTCGCCAGAGAAATATCCGAGCGGCTCCGCGACGACAGAGAGTGGGCG 76
QY	75	GGATGTGTGTGGCGGGCTGCGGCTCCTCGTCCCTCCAGCGGCGCTGAGCGGCATCGA 134
DB	77	GGATGTGTGTGGCGGGCTGCGGCTCCTCGTCCCTCCAGCGGCGCTGAGCGGCATCGA 136
QY	135	TTTGTCCTGGGCGGCGACCGTCGACCCGCGGAGATGAGGCGCTGATTGCAAGTTAA 194
DB	137	TTTGTCCTGGGCGGCGACCGTCGACCCGCGGAGATGAGGCGCTGATTGCAAGTTAA 196
QY	195	AAGTAAAGAACCAATGCTCAATTTTCCAAACCTTTTGGTGGACCTCGATATCTGGGC 254
DB	197	AAGTAAAGAACCAATGCTCAATTTTCCAAACCTTTTGGTGGACCTCGATATCTGGGC 256
QY	255	CATACTGTAGAGAAAGAGCAAGCATGATCAGAGTTCATGTTTAAAGCCAAATATC 314
DB	257	CATACTGTAGAGAAAGAGCAAGCATGATCAGAGTTCATGTTTAAAGCCAAATATC 316
QY	315	TGATTCATTACTGGTGTATCAAGCTAGAAACTTTTTCATCTGGGTACCTCAAC 374
DB	317	TGATTCATTACTGGTGTATCAAGCTAGAAACTTTTTCATCTGGGTACCTCAAC 376
QY	375	TGTTTACACAGATATGGGCACTTAGCTGACATGATATATGATGAGAGATGATCAAGT 434
DB	377	TGTTTACACAGATATGGGCACTTAGCTGACATGATATATGATGAGAGATGATCAAGT 436
QY	435	GGAGTTTCCATAGTATGAACCTATCAACTGAAGGTACAGAAATTCAGTACCCTC 494
DB	437	GGAGTTTCCATAGTATGAACCTATCAACTGAAGGTACAGAAATTCAGTACCCTC 496
QY	495	TGCACT 500
DB	497	TGCACT 502
RESULT 2		
LOCUS	BQ050397	1089 bp mRNA linear EST 29-MAR-2002
DEFINITION	AGNCOURT 7050803 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5784342	
ACCESSION	BQ050397	5' mRNA sequence.
VERSION	BQ050397.1	GI:19809737
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1 (bases 1 to 1089)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contract: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LML at:
<http://image.llnl.gov>
 Plate: LML12869 row: 1 column: 07
 High quality sequence stop: 612.
 Location/Qualifiers
 1..1089
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5784342"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NH_MGC_71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."
 ORIGIN
 Query Match 97.2%; Score 486; DB 12; Length 1089;
 Best Local Similarity 100.0%; Pred. No. 77e-116;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 15 GTACGGCGGCTCGCAGGAGAATCCGAGGGGCTCGGGACGACAGAGGGCGGGC 74
 Db 57 GTACGGCGGCTCGCAGGAGAATCCGAGGGGCTCGGGACGACAGAGGGCGGGC 116
 Oy 75 GGGATGATGTGCGGGGCTCGGGCTCTCTGCTCCCGACGGCGCGCTGACGGGCACTGA 134
 Db 117 GGGATGATGTGCGGGGCTCGGGCTCTCTGCTCCCGACGGCGCGCTGACGGGCACTGA 176
 Oy 135 TTGTGCTCGGGGCGGCGAGCGCGACCCCGCGAGATGAGGCGTGCATTAGCAAGTTAA 194
 Db 177 TTGTGCTCGGGGCGGCGAGCGCGACCCCGCGAGATGAGGCGTGCATTAGCAAGTTAA 236
 Oy 195 AAGTACGAGAACATGGCTCAGTTTCCAAACCTTTGGTGGCGAGCCGTGATCTGGGC 254
 Db 237 AAGTACGAGAACATGGCTCAGTTTCCAAACCTTTGGTGGCGAGCCGTGATCTGGGC 296
 Oy 255 CATACGTGTAGAGGAAAGCGAAGCATGATCAGAGTTCATAGTTTAAAGCCAAATATC 314
 Db 297 CATACGTGTAGAGGAAAGCGAAGCATGATCAGAGTTCATAGTTTAAAGCCAAATATC 356
 Oy 315 TGGATTCTATCTGGTGATCAAGCTAGAACTTTTTTCAATCTGGGTTAACTCAAC 374
 Db 357 TGGATTCTATCTGGTGATCAAGCTAGAACTTTTTTCAATCTGGGTTAACTCAAC 416
 Oy 375 TGTTTTACACAGATATGGGACCTAGAGTCGATGATATATGAGAAATGATGCAAGT 434
 Db 417 TGTTTTACACAGATATGGGACCTAGAGTCGATGATATATGAGAAATGATGCAAGT 476
 Oy 435 GGAAGTTTCCATAGCTATGAAACTTATCAACTGAAAGCTACAGAGTATCGACTACCTC 494
 Db 477 GGAAGTTTCCATAGCTATGAAACTTATCAACTGAAAGCTACAGAGTATCGACTACCTC 536
 Oy 495 TGCACCT 500
 Db 537 TGCACCT 542
 RESULT 3
 LOCUS B0941411 979 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGNCOURT 8744157 Lupski sciatic nerve Homo sapiens CDNA clone

IMAGE:6205905 5', mRNA sequence.
ACCESSION B0941411
VERSION B0941411.1 GI:22356889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 979)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLM13629 row: 1 column: 10
High quality sequence stop: 599.
Location/Qualifiers
1..979
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6205905"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; CDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCCAGGCTCCG-3' and
5'-GACTAGTTCATGATCGAGCGGCGCCCTT(15)-3'. Site selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
ORIGIN
Query Match 96.9%; Score 484.4; DB 13; Length 979;
Best Local Similarity 99.8%; Pred. No. 2e-115;
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
15 GTACGGGGCTCGCGAGAGAAATCCGAGCGGGCTCCGGGACGAGACAGAGGGGGCG 74
Db 93 GTACGGGGCTCGCGAGAGAAATCCGAGCGGGCTCCGGGACGAGACAGAGGGGGCG 152
Qy 75 GGGATGATGTCGGGGCTCGGGCTCCTGCTCCGAGCGGGCGTGAAGGGGACATGA 134
Db 153 GGGATGATGTCGGGGCTCGGGCTCCTGCTCCGAGCGGGCGTGAAGGGGACATGA 212
Qy 135 TTTGTCCCTGGGGCGGACCGCGGACCCGCCGAGATGAGCGCTGATTAGCAAGTAA 194
Db 213 TTTGTCCCTGGGGCGGACCGCGGACCCGCCGAGATGAGCGCTGATTAGCAAGTAA 272
Qy 195 AAGTAAACAGAACATGCTGCTGATTTTCAACACCTTTTGGTGGAGCTGATATCTGGGC 254
Db 273 AAGTAAACAGAACATGCTGCTGATTTTCAACACCTTTTGGTGGAGCTGATATCTGGGC 332
Qy 255 CATACCTGAGAGGAAGAGGAGAGCATGATGAGAGTTTCAAGTTTAAACCAATATC 314
Db 333 CATACCTGAGAGGAAGAGGAGAGCATGATGAGAGTTTCAAGTTTAAACCAATATC 392
Qy 315 TGGATTCACTTACTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTACCTCAAC 374
Db 393 TGGATTCACTTACTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTACCTCAAC 452

Qy 375 TGGTTAGCAGATATGGGACATGATGATGATGATGATGATGATGATGATGATGAT 434
Db 453 TGGTTAGCAGATATGGGACATGATGATGATGATGATGATGATGATGATGATGAT 512
Qy 435 GGAGTTTCCATAGCTATGAACTTATCAACTGAGCTAGACAGATATCAAGTACCTC 494
Db 513 GGAGTTTCCATAGCTATGAACTTATCAACTGAGCTAGACAGATATCAAGTACCTC 572
Qy 495 TGCACCT 500
Db 573 TGCACCT 578
RESULT 4
Bg118422 750 bp mRNA linear EST 30-JAN-2001
LOCUS 602348040P1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:444129 5',
DEFINITION mRNA sequence.
ACCESSION Bg118422 GI:12611928
VERSION Bg118422.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 750)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLM10216 row: p column: 10
High quality sequence stop: 642.
Location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:444129"
/tissue_type="adenocarcinoma, cell line"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 94.8%; Score 474; DB 12; Length 750;
Best Local Similarity 99.8%; Pred. No. 9.4e-113;
Matches 485; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
15 GTACGGGGCTCGCGAGAGAAATCCGAGCGGGCTCCGGGACGAGACAGAGGGGGCG 74
Db 101 GTACGGGGCTCGCGAGAGAAATCCGAGCGGGCTCCGGGACGAGACAGAGGGGGCG 160
Qy 75 GGGATGATGTCGGGGCTCGGGCTCCTGCTCCGAGCGGGCGTGAAGGGGACATGA 134
Db 161 GGGATGATGTCGGGGCTCGGGCTCCTGCTCCGAGCGGGCGTGAAGGGGACATGA 220
Qy 135 TTTGTCCCTGGGGCGGACCGCGGACCCGCCGAGATGAGCGCTGATTAGCAAGTAA 194
Db 221 TTTGTCCCTGGGGCGGACCGCGGACCCGCCGAGATGAGCGCTGATTAGCAAGTAA 280

QY	195	AAGTAAACGAAACATAGCTCAGATTCTCCAAACACTTTTGATGGACCTCGATATCTGGGC	254
Db	281	AAGTAAACGAAACATAGCTCAGATTCTCCAAACACTTTTGATGGACCTCGATATCTGGGC	340
QY	255	CATACACTGTAGAGAAAGAGCTCAGATGATCAGCAGTTTCCATAGTTTAAAGCCAAATATC	314
Db	341	CATACACTGTAGAGAAAGAGCTCAGATGATCAGCAGTTTCCATAGTTTAAAGCCAAATATC	400
QY	315	TGGAATTCATTACTGGTGTATCAAGCTAGAAACTTTTTCATCTGGGTTACTCTCAAC	374
Db	401	TGGAATTCATTACTGGTGTATCAAGCTAGAAACTTTTTCATCTGGGTTACTCTCAAC	460
QY	375	TGTTTACCAAGAAATAGGCGCATAGCTGCACATGAATATATATGAAAGATAGATCAAGT	434
Db	461	TGTTTACCAAGAAATAGGCGCATAGCTGCACATGAATATATGAAAGATAGATCAAGT	520
QY	435	GGAGTTTCCATAGCTATGAATCTTATCAAACTGAAGCTACAGGATATCAGTCAACCTC	494
Db	521	GGAGTTTCCATAGCTATGAATCTTATCAAACTGAAGCTACAGGATATCAG-TACCTC	579
QY	495	TGCACT 500	
Db	580	TGCACT 585	

RESULT 5	BC020269	2126 bp	linear	HTC 19-DEC-2001
LOCUS	BC020269			
DEFINITION	Homo sapiens, clone IMAGE:4898011, mRNA.			
ACCESSION	BC020269			
VERSION	BC020269.1			
KEYWORDS	HTC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (19-DEC-2001)
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
1 (bases 1 to 2126)
Strausberg, R
Direct Submission
National Institutes of Health, Mammalian

REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: CGAPbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Robin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

Steven Joneau, Jennifer Asano, Ian Bosdel, Aaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Gardland, Ran Guin, Letlicia Hsiao, Martin Krzyzanski, Reta Kutischek, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven Ness, Pawan Pandoh, Anna-Liisa Pihlbu, Parvaneh Seedei, Jacqueline Schinn, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stoltz, Michael Thorne, Miranad Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuydamdadi, Marco Marra.

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN/ at: <http://image.llnl.gov>
Series: IRBL plate: 40 row: n Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796
This clone has the following problem: no cloning site /
microdeletion.

FEATURES
source

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"
```

```

/clone="IMAGE:4899011"
/tissue_type="Pancreas, epithelioid carcinoma"
/clone_id="NH MG 42"
/lab_host="DH10B-R-"
/note="Vector: pOTB7"

```

Query Match	94.8%;	Score 474;	DB 11;	Length 2126;
Best Local Similarity	100.0%;	Pred. No. 1.2e-112;		
Matches 474;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	27	GCAGAGGAAGATCCCGAGCGGGCTCCGGGACGGACAGAGAGCGCGGGGATGCGTGTG	86
Db	1	GCAGAGGAAGATCCCGAGCGGGCTCCGGGACGGACAGAGAGCGCGGGGATGCGTGTG	60
QY	87	GGGGCTGGGGCTCTCGGCTCCCTCCAGCGGCGTGAAGCGGACTGATTTGCTCTGGG	146
Db	61	GGGGCTGGGGCTCTCGGCTCCCTCCAGCGGCGCTGAAGCGGCACTGATTTGCTCTGGG	120
QY	147	GGCGCAGGGGGGACCCCGCGGGAGATGAGGGGTGATTAAGCAAGGTAAAGTAACAAGAC	206
Db	121	GGCGCAGGGGGGACCCCGCGGGAGTGAAGGGGTGATTAAGCAAGGTAAAGTAACAAGAC	180
QY	207	CATGGCTCAGTTTCCACACCTTTTGGTGGAGCCTCGATATCTGGGCCATAACTGTAGA	266
Db	181	CATGGCTCAGTTTCCACACCTTTTGGTGGAGCCTCGATATCTGGGCCATAACTGTAGA	240
QY	267	GGAAGAGCCGACGATGATCAGCAGTTTCATAGTTTAAAGCCAAATCTGGAATCATTC	326
Db	241	GGAAGAGCCGACGATGATCAGCAGTTTCATAGTTTAAAGCCAAATCTGGAATCATTC	300
QY	327	TGGTGATCAAGCTCAGAACTTTTTCCTCAATCTGGGTACCTCAACCTGTTTAGACA	386
Db	301	TGGTGATCAAGCTCAGAACTTTTTCCTCAATCTGGGTACCTCAACCTGTTTAGACA	360
QY	387	GATATGGGCATCTAGCTGACATGAATATATGATGGAAGAAATGATCAAGTGGAGTTTTCAT	446
Db	361	GATATGGGCATCTAGCTGACATGAATATATGATGGAAGAAATGATCAAGTGGAGTTTTCAT	420
QY	447	AGCATGAACTTATCAACTGAAGCTCAGAGGATATAGCTCAAGCCTCTGCACT	500
Db	421	AGCATGAACTTATCAACTGAAGCTCAGAGGATATAGCTCAAGCCTCTGCACT	474

RESULT 6	EX470886	643 bp	mRNA	linear	EST 04-SEP-2003
LOCUS	EX470886				
DEFINITION	DEF2686617123_r1 686 (synonym: hlcc3) Homo sapiens CDNA clone DREF2686617123_5', mRNA sequence.				

VERSION	BX470886.1	GI:31665209
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	

ORGANISM	Homio sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 643)
REFERENCE	Bahr,A., Lauber,J., Mewes,H.W., Well,B., Amlid,C., Osanger,A., Rebo,G., Han,M. and Wiemann,S. EST (Bahr,A., Lauber,J., Mewes,H.W., Well,B., et al.) Unpublished (2003)
JOURNAL	Contact: MIPS
COMMENT	

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert from S. Wiemann, Molecular Genome Analysis, Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project. No 3' sequence available.

This clone (DKFZp686E17123) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
SOURCE
Location/Qualifiers
1. .643
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686E17123"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hicc3)"
/note="Vector: pTriplex2; Site_1: SfilA; Site_2: SfilB;
cDNA-collection"

ORIGIN
Query Match 92.0%; Score 460; DB 13; Length 643;
Best Local Similarity 99.8%; Pred. No. 4.1e-109;
Matches 471; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
29 GAGGAAGATCCGAGCGGGCTCCGGAGCGAGACAGAGGGGGGAGTGTGTGG 88
2 GAGGAGAGATCCGAGCGGGCTCCGGAGCGAGACAGAGGGGGGAGTGTGTGG 61
89 GGCTGCGGCTCTGCTGCTCTCCAGCGCGCGGTGAGCGGCACTGATTGTCTCGGGC 148
62 GGCTGCGGCTCTGCTGCTCTCCAGCGCGCGGTGAGCGGCACTGATTGTCTCGGGC 121
149 GCGAGCGGAGCCGCGCGGAGATGAGCGCTGATTAACAAGTTAAAGTAAAGAACCA 208
122 GCGAGCGGAGCCGCGCGGAGATGAGCGCTGATTAACAAGTTAAAGTAAAGAACCA 181
209 TGGCTCAATTCGAACACCTTTGGTGGAGCGCTGAGATCTGGGCCATTACTGAGAG 268
182 TGGCTCAATTCGAACACCTTTGGTGGAGCGCTGAGATCTGGGCCATTACTGAGAG 241
269 AAAGAGCGAGATGATGAGAGATTCATAGTTTAAAGCAATCTGGATTCTACTG 328
242 AAAGAGCGAGATGATGAGAGATTCATAGTTTAAAGCAATCTGGATTCTACTG 301
329 GTGATCAAGTGAACCTTTTTCATCTGGGTTACCTGACCTGTTTGAACAGCA 388
302 GTGATCAAGTGAACCTTTTTCATCTGGGTTACCTGACCTGTTTGAACAGCA 360
389 TATGGGCACTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 448
361 TATGGGCACTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
449 CTATGAACCTATCAAACTGAAGCTACAAGATATCACTACCTCTGCACT 500
421 CTATGAACCTATCAAACTGAAGCTACAAGATATCACTACCTCTGCACT 472

RESULT 7
LOCUS BG829540 811 bp mRNA linear EST 22-MAY-2001
DEFINITION 602763842P1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4899011 5',
mRNA sequence.
ACCESSION BG829540
VERSION BG829540.1 GI:14177114
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strussberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/ULNL at:
http://image.jml.gov
Plate: L10M1790 row: c column: 12
High quality sequence stop: 651.
Location/Qualifiers
1. .811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4899011"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 92.0%; Score 459.8; DB 12; Length 811;
Best Local Similarity 99.4%; Pred. No. 4.9e-109;
Matches 472; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
27 GCGAGGAAGATCCCGAGCGGCTCCGGAGCGAGACAGAGCGCGGAGTGTGTGC 86
2 GCGAGGAAGATCCCGAGCGGCTCCGGAGCGAGACAGAGCGCGGAGTGTGTGC 61
87 GGGGCTGGGCTCTGCTGCTCTCCAGCGGCGCGTGAAGCGGCACTGATTGTCTCTGG 146
62 GGGGCTGGGCTCTGCTGCTCTCCAGCGGCGCGTGAAGCGGCACTGATTGTCTCTGG 121
147 GCGGAGCGGCGGAGCCGCGGAGATGAGGCGTGTGATGAGAGGTAAGTAAGACAGAAC 206
122 GCGGAGCGGCGGAGCCGCGGAGATGAGGCGTGTGATGAGAGGTAAGTAAGACAGAAC 181
207 CATGGCTCAGTTTCCAAACCTTTTGGTG-GCAGCTGTGATATCTGGGCCATTACTGTAG 265
182 CATGGCTCAGTTTCCAAACCTTTTGGTGCGCAGCTGTGATATCTGGGCCATTACTGTAG 241
266 AGGAAGAGCGAGATGATGAGAGATTCATAGTTTAAAGCAATATCTGGATTCTACTG 325
242 AGGAAGAGCGAGATGATGAGAGATTCATAGTTTAAAGCAATATCTGGATTCTACTG 301
326 CTGGGATCAAGCTGAGAACTTTTTCATCTGGGTTACCTCAACCTGTTTAGGAC 385
302 CTGGGATCAAGCTGAGAACTTTTTCATCTGGGTTACCTCAACCTGTTTAGGAC 361
386 AGATATGGGCACTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 445
362 AGATATGGGCACTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
446 TAGCTATGAACCTATCAAACTGAAGCTACAAGATATCACTACCTCTGCACT 500
422 TAGCTATGAACCTATCAAACTGAAGCTACAAGATATCACTACCTCTGCACT 476

RESULT 8
LOCUS AL711737 544 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686K1884.r1 686 (synonym: hicc3) Homo sapiens cDNA clone
DKFZp686K1884 5', mRNA sequence.
ACCESSION AL711737
VERSION AL711737.1 GI:19695092
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 544)
 AUTHORS Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 TITLE EST (Duesterhoeft, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: MIPS

MIPS
 Ingolstaedter Landstr.1, D-85764 Neuberg, Germany
 This is the 5' sequence of the clone insert
 clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No 5' sequence available.
 This clone (DKFZ66861884) is available at the RZPD in Berlin.
 Please contact the RZPD; Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source

1..544
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZ66861884"
 /dev_stage="adult"
 /lab_host="DH1B"
 /note="Vector: pTRipleX2; Site_1: SfIIA; Site_2: SfIIB;
 cDNA-collection"

ORIGIN

Query Match 90.8%; Score 454; DB 9; Length 544;
 Best Local Similarity 99.8%; Pfd. No. 1.4e-107;
 Matches 465; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

35 GATCCGAGCGGCGCTCCGCGACGACAGAGAGCGCGGGATGCTGCGGGCTGC 94
 |||||
 2 GATCCCGAGCGGCGCTCCGCGACGACAGAGAGCGCGGGATGCTGCGGGCTGC 61
 |||||
 95 GGCTCCTGCGCTCCGCGACGAGCGGTGAGCGGCGGCGGCGGCGGCGGCGG 154
 |||||
 62 GGCTCCTGCGCTCCGCGACGAGCGGTGAGCGGCGGCGGCGGCGGCGGCGG 120
 |||||
 155 GCGGACCGCGCGGAGATGAGCGGTGATTAGCAAGGTAAAGTAAAGTAAAGT 214
 |||||
 121 GCGGACCGCGCGGAGATGAGCGGTGATTAGCAAGGTAAAGTAAAGTAAAGT 180
 |||||
 215 AGTTTCACACCTTTTGGTGGAGCGCTGATCTGGGCGCATACGTAGAGAGAG 274
 |||||
 181 AGTTTCACACCTTTTGGTGGAGCGCTGATCTGGGCGCATACGTAGAGAGAG 240
 |||||
 275 CGAAGCATGATGACGAGTTCATAGTTTAAAGCCAAATCTGGATTCATTACTG 334
 |||||
 241 CGAAGCATGATGACGAGTTCATAGTTTAAAGCCAAATCTGGATTCATTACTG 300
 |||||
 335 AGCTTGAACCTTTTTCATCTGGGTTACCTCAACCTGTTTGAAGCAGATATG 394
 |||||
 301 AAGCTAAGAACTTTTTCATCTGGGTTACCTCAACCTGTTTGAAGCAGATATG 360
 |||||
 395 CACTAGCTGACATGATATGATGAAAGATGATCAAGTGAGTTTCATAGCATGA 454
 |||||
 361 CACTAGCTGACATGATATGATGAAAGATGATCAAGTGAGTTTCATAGCATGA 420
 |||||
 455 AACTATCAACTGAAGCTACAAAGATATCAAGTACCTCTGCACT 500
 |||||
 421 AACTATCAACTGAAGCTACAAAGATATCAAGTACCTCTGCACT 466
 |||||

RESULT 9
 CB296595 605 bp mRNA linear EST 28-FEB-2003
 LOCUS 12B22043 rev 1 A05 r 037.abi Chimpanzee brain library Koo's Pan
 DEFINITION troglodytes cDNA clone 12B22043_rev_1_A05_r_037.abi 5', mRNA
 sequence.

ACCESSION CB296595
 VERSION CB296595.1 GI:28622025
 KEYWORDS EST.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes

REFERENCE 1 (bases 1 to 605)
 AUTHORS Hellmann, I., Zollner, S., Enard, W., Ebersberger, I., Nickel, B. and Paabo, S.
 TITLE Selection on human genes as revealed by comparisons to chimpanzee
 JOURNAL Genome Res. (2003) In press

COMMENT Contact: Paabo S
 Evolutionary Genetics
 Max-Planck-Institute for evolutionary Anthropology
 Deutscher Platz 6, 04103 Leipzig, Germany
 Tel: +49-(0)-341-3550 500
 Fax: +49-(0)-341-3550 555
 Email: paabo@eva.mpg.de
 Seq primer: M13 reverse.

FEATURES
 source

1..605
 /organism="Pan troglodytes"
 /mol_type="mRNA"
 /db_xref="taxon:9598"
 /clone="12B22043_rev_1_A05_r_037.abi"
 /sex="male"
 /tissue_type="brain, presumably cortex"
 /dev_stage="adult"
 /lab_host="Epicurian Coli (TM) Xu-10-Gold"
 /clone_lib="Chimpanzee brain library Koo's"
 /note="Vector: pUC19; Site_1: SfII-A; Site_2: SfII-B; The
 library was prepared using the SMART cDNA library
 construction kit (Clontech), doing only primer extension,
 but not PCR amplification of the cDNA. The only deviation
 from the published protocol was that we cloned the cDNA
 into a plasmid vector."

ORIGIN

Query Match 87.5%; Score 437.4; DB 14; Length 605;
 Best Local Similarity 99.6%; Pfd. No. 3.2e-103;
 Matches 449; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

50 TCCGGGACGACAGAGAGCGGCGGAGTGGTGGGCGGCGGCGGCGGCGGCGG 109
 |||||
 1 TCCGGGACGACAGAGAGCGGCGGAGTGGTGGGCGGCGGCGGCGGCGGCGG 60
 |||||
 110 CCCAGCGGCGGCGTGAAGCGGCACTGATTGTCCCTGGGCGGCGGCGGCGG 169
 |||||
 61 CCCAGCGGCGGCGTGAAGCGGCACTGATTGTCCCTGGGCGGCGGCGGCGG 120
 |||||
 170 GATGAGCGCTGATGACAGAGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 229
 |||||
 121 GATGAGCGCTGATGACAGAGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 180
 |||||
 230 TTGGTGGCAGCTGATATCTGGGCCATTAAGTAAAGTAAAGTAAAGTAAAGT 289
 |||||
 181 TTGGTGGCAGCTGATATCTGGGCCATTAAGTAAAGTAAAGTAAAGTAAAGT 240
 |||||
 290 AGTTTCATGTTTAAAGCCAAATCTGATTTACTGTTGATCAAGTAAAGTAA 349
 |||||
 241 AGTTTCATGTTTAAAGCCAAATCTGATTTACTGTTGATCAAGTAAAGTAA 299
 |||||
 350 TTTTTCATCTGGTTACTCAACCGTTTGGACAGATATGGGAGCTAGCTGACAT 409
 |||||
 300 TTTTTCATCTGGTTACTCAACCGTTTGGACAGATATGGGAGCTAGCTGACAT 359
 |||||
 410 ATATATGATGAAGATGATCAAGTGAAGTTTTCATAGCTATGAAGTAACTTA 469
 |||||
 360 ATATATGATGAAGATGATCAAGTGAAGTTTTCATAGCTATGAAGTAACTTA 419
 |||||
 470 AGCTACAGAGATATCAAGTACCTCTGCACT 500
 |||||

```

DB      420 AGCTAAGAGATATCAGCTACCCCTGCACT 450
|||||
RESULT 10
BC013578
LOCUS   Homo sapiens, Similar to Intersectin 1 (SH3 domain protein), clone
DEFINITION
ACCESSION BC013578      2079 bp      mRNA      linear      HTC 04-MAR-2003
IMAGE:3878242, mRNA.
VERSION   BC013578.1   GI:15488896
KEYWORDS  HTC.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2079)
AUTHORS   Strausberg, R.
TITLE     Direct Submission
JOURNAL   Submitted (04-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNI)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LILNI at: http://image.llnl.gov
Series: IRAX Plate: 14 Row: C Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504796
This clone has the following problem: retained intron.
FEATURES
Source
1..2079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="3878242"
/tissue_type="lung, large cell carcinoma"
/clone_lib="NIH MGC_68"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
ORIGIN
Query Match      84.8%; Score 424; DB 11; Length 2079;
Best Local Similarity 100.0%; Pred. No. 1.4e-99;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
77 GATGCTGTCGGGGCTGCGGCTCTGCGTCTCCACAGGGGCGCTGAGCGGCACTGATT 136
DB      1 GATGCTGTCGGGGCTGCGGCTCTGCGTCTCCACAGGGGCGCTGAGCGGCACTGATT 60
137 TGTCCCTGGGGCGGACGCGGACCCGCCGAGATGAGGCGTGATTAGCAAGGTAAAA 196
DB      61 TGTCCCTGGGGCGGACGCGGACCCGCCGAGATGAGGCGTGATTAGCAAGGTAAAA 120
197 GTAAGAGACCATGGCTGATTTTCAACACCTTTTGGTGGAGCGCTGATATCTGGGCA 256
DB      121 GTAAGAGACCATGGCTGATTTTCAACACCTTTTGGTGGAGCGCTGATATCTGGGCA 180
257 TAACTGTAGAGAGAGAGAGAGATGATCAGACAGTTCATAGTTAAAGCCATATCTG 316
DB      181 TAACTGTAGAGAGAGAGAGAGATGATCAGACAGTTCATAGTTAAAGCCATATCTG 240

```

```

QY      317 GATTCATTACTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACTCAACTGG 376
DB      241 GATTCATTACTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACTCAACTGG 300
QY      377 TTTTAGACACAGATATNGGGCACTAGCTGACATGAATTAAGTGAAGATGATCAAGTGG 436
DB      301 TTTTAGACACAGATATNGGGCACTAGCTGACATGAATTAAGTGAAGATGATCAAGTGG 360
QY      437 AGTTTCCATAGCTATGAACTTATCAAACTGAAGCTACAAGATATGACCTCCCTGG 496
DB      361 AGTTTCCATAGCTATGAACTTATCAAACTGAAGCTACAAGATATGACCTCCCTGG 420
QY      497 CACT 500
DB      421 CACT 424
|||||
RESULT 11
CB269754
LOCUS   CB269754      734 bp      mRNA      linear      EST 20-FEB-2003
DEFINITION
ACCESSION CB269754      Human Fat Cell 5'-Stretch Plus CDNA Library Homo sapiens
CDNA 5', mRNA sequence.
VERSION   CB269754.1   GI:28444339
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 734)
AUTHORS   Yang, R.-Z., Shuldiner, A. and Gong, D.-W.
TITLE     EST analysis of human adipose gene expression
JOURNAL   Unpublished (2002)
Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, BH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR Primers
FORWARD: CTCGGGAAGCGCGCATTTGTGTGT
BACKWARD: AATACGACTCATATGAGGCGCAATTGG
Seq primer: GTTGCTACCCGGGAATTC.
FEATURES
Source
1..734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus CDNA Library"
/note="Vector: lambda triplex"
ORIGIN
Query Match      82.6%; Score 413; DB 14; Length 734;
Best Local Similarity 100.0%; Pred. No. 7.8e-97;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
88 GGGCTGGGGCTCTCCGCGTCCCTCCACAGCGGCGGTGACGACAGATTTGGCCCTGGGG 147
DB      1 GGGCTGGGGCTCTCCGCGTCCCTCCACAGCGGCGGTGACGACAGATTTGGCCCTGGGG 60
148 CGGACGCGGACCGCGCGGAGATGAGCGGTGATTAGCAAGTAAAGTAACAGAAC 207
DB      61 CGGACGCGGACCGCGCGGAGATGAGCGGTGATTAGCAAGTAAAGTAACAGAAC 120
208 ATGGCTCAGTTTCAACACCTTTTGGTGGAGCGCTGATATCTGGGCCATTACTGTAGG 267
DB      121 ATGGCTCAGTTTCAACACCTTTTGGTGGAGCGCTGATATCTGGGCCATTACTGTAGG 180
268 GAAAGAGGAGAGATGATCAGACAGTTCATAGTTAAAGCCATATCTGATTAATCT 327

```

D _b	181	GAAAGAGGGAAGCATGATCAGTACGTTCATGTTTAAAGCCAAATCTGGATTCAACT	240
O _y	328	GGTGATCAGGTAGAAACTTTTTTTTCAATCGGGTTAACCCGAACGTTTACACAAG	387
D _b	241	GGTAGTCAGGTAGAAACTTTTTTTTTCAATCTGGGTTAACCCTCAACCTGTTTTACACAAG	300
O _y	388	ATTATGGGCATCAGCTGACATGATAATATGATGGAAACATGGATCAAGTGGAGTTTCCATA	447
D _b	301	ATTATGGGCATCAGCTGACATGATAATATGATGGAAACATGGATCAAGTGGAGTTTCCATA	360
O _y	448	GCTATGAACCTTATCAAACCTGTAAGGTACAAGGATATCAACCTCTGCACCT	500
D _b	361	GCTATGAACCTTATCAAACCTGTAAGGTACAAGGATATCAACCTCTGCACCT	413

RESULT	12
A1750406	
LOCUS	A1750406
DEFINITION	405 bp mRNA linear EST 20-JUN-2007 cn02g92.x2 Normal Human trabecular Bone Cells Homo sapiens CDNA clone NHTBC_cn02g02 random_mRNA sequence.
ACCESSION	A1750406
VERSION	A1750406.1 GI:5128670
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 405)
AUTHORS	Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G.,

JOURNAL COMMENT	Unpublished (1997) Contact: Libin Jia

Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).
plate: 02 row: g column: 02
Seq primer: -21m3 forward primer (ABI).
Location/Qualifiers
1. .405

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NHTRC_fm02902"
/sex="Female"
/tissue_type="bone"
/cell_type="Tribecular Bone Cells"
/lab_host="SUPR"
/clone_1ib="Normal Human Trabecular Bone Cells"
/notes="Organ: Hip; Vector: pBluescript; Site 1: EcoRI;
Library constructed by Dr. Marian Young and Dr. Pamela
Georion Robey (NIDCR)"

```

Query March	79.5%	Score 397.6	DB 9	Length 405
Best Local Similarity	99.0%	Pied. No. 6.9e-93		
Matches 400; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

Db 2 GCACGAGCCGGAGCAGACAGACAGCGCGGCGGGGATGTTGTCGGGGCTGCGGCTCCAGC 61

QY 104 GTCCCTCCAGCGGCGGTAGCGGCACTGATTTGTCCTGGGGCGGAGCGGACCCG 163
 |||||
 Db 62 GTCCTCCAGCGGCGGTAGCGGCACTGATTTGTCCTGGGGCGGAGCGGACCCG 122

OY	164	CCCCGAGATGAGGCGTCGATTGCAAGGTTAAAAAGTACAGAACCATGCTCAAGTTTCCAA	223
Db	122	CCCCGAGATGAGGCGCTCCATTGCAAGGTTAAAAAGTACAGAACCATGCTCAAGTTTCCAA	181
OY	224	CACCTTTGGTGGCGACCTGGAATCTGGGCGCATTAAGGAAAGAGGAAAGCATG	283
Db	182	CACCTTTGGTGGCGACCTGGATATCTGGGCGCATTAAGGAAAGAGGAAAGCATG	241
OY	284	ATCAGCACTTCCATGTTTAAAGCCAAATATCTGGAATCAATTAAGTGTATCAAGCTAGAA	343
Db	242	ATCAGCACTTCCATGTTTAAAGCCAAATATCTGGAATCAATTAAGTGTATCAAGCTAGAA	301
OY	344	ACTTTTTTTTCAATCTGGTTTAACTCTCAACCTGTTTATGACAGATATGGGCACTAGCTG	403
Db	302	ACTTTTTTTTCAATCTGGTTTAACTCTCAACCTGTTTATGACAGATATGGGCACTAGCTG	361
OY	404	ACATGAATTAATGATGGAGAAATGATCAAGTGAAGTTTCCATA	447
Db	362	ACATGAATTAATGATGGAGAAATGATCAAGTGAAGTTTCCATA	405

RESULT	13
Locus	Bf194430
Definition	Bf194430 547 bp mRNA linear EST 02-NOV-2000
Accession	Z6586 MARCH 2 PIG Sus scrofa CDNA 5' mRNA sequence.
Version	Bf194430
Keywords	Bf194430.1 GI:11077799
Source	EST.
Organism	Sus scrofa (pig)
	Sus scrofa

REFERENCE
AUTHORS
1 (bases 1 to 547)
Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
Vallet, J., Wise, T., Rohrer, G.A., Perte, G., Sultana, R.,

TITLE	Quackenbush, J. and Keele, J.W. Pooled gene discovery by normalized cDNA-library sequencing and Erm cluster assembly
JOURNAL	Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE	2213789
PUBMED	12226715
COMMENT	Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smilh@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
-and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 76 row: E column: 21
Seq primer: ATTAGGTGACACTATAG.

```

FEATURES
source
location/Qualifiers
1..547
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="PDH10p"
/clone_idb="MARC 2P1G"
/note="vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

```

Query Match	77.5%;	Score 387.6;	DB 10;	Length 547;
Best Local Similarity	93.3%;	Pred. No. 3e-90;		
Matches 405, Conservative	0;	Mismatches 29;	Indels 0;	Gaps 0;
Oy	67	GGCGGAGCGGAGATGCTTCGCGGGCTCCGCTCCCTCCAGCGGCGGTAGC	126	

```
Db      2 GGGGGGGGGGATGCTGTGCAAGGCTGGGCTCTCGCTCTCTCTCGGGGAGTGAAC 61
QY      127 GGCACGTATTTGTCCTTGGGGGGGAGCGCGGACCCGCCGAGATGAGCGTCGATTAG 186
Db      62 TGCACTGATTTGTCTGTGGGGGGGAGCGCGGACCCGCCGAGATGAGCGTCGATTAG 121
QY      187 CAAGTAAAGTAAAGAAACCAATGCTGCTGATTTCCAAACCTTTGGTGGCAGCTGAT 246
Db      122 CAAGGTAAAGTAAAGAAACCAATGCTGCTGATTTCCAAACCTTTGGTGGCAGCTGAT 181
QY      247 ATCTGGGCCATTAAGTGTAGAGAAAGAGAGAGATGATGATGATGATGATGATGAT 306
Db      182 ATTTGGGCCATTAAGTGTAGAGAAAGAGAGAGATGATGATGATGATGATGATGAT 241
QY      307 CCAATATCTGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 366
Db      242 CCAATATCTGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
QY      367 CCGTAACTGTTTATGACAGATATGAGCAGTATGATGATGATGATGATGATGATGATGAT 426
Db      302 CCTCAACTGTTTATGACAGATATGAGCAGTATGATGATGATGATGATGATGATGATGAT 361
QY      427 GATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTGAACTGAACTGAACT 486
Db      362 GATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTGAACTGAACTGAACT 421
QY      487 CTACCCCTTGCACT 500
Db      422 CTCCCTCCGCACT 435

RESULT 14
CB519152      698 bp      mRNA      linear      EST 09-JUL-2003
LOCUS      UI-M-GH0-ceh-c-14-0-UI.r1 NIH BMAP_GH0 Mus musculus cDNA clone
DEFINITION      IMAGE:6839463 5', mRNA sequence.
ACCESSION      CB519152.1 GI:29352507
VERSION      EST.
KEYWORDS      Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS      1 (bases 1 to 698)
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
FEATURES
source
location/Qualifiers
1..698
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6839463"
/tissue_type="whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (TI phage resistant)"
/clone_id="NIH BMAP GH0"
/note="Organ: Brain; Vector: pYX-Asc; Site_1: Ecor I;
```

```
Site 2: Not 1: The library was constructed according
Bonaïdo, Lemon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator.'

ORIGIN
Query Match      73.5%; Score 367.4; DB 14; Length 698;
Best Local Similarity 85.5%; Pred. No. 6.1e-85;
Matches 425; Conservative 0; Mismatches 61; Indels 11; Gaps 1;

QY      15 GTACGGCGGCTCGGAGGAGAAATCCGAGCGGGCTCCGGAGCGA-----CAG 63
Db      58 GTACGGCGGCTCGGAGGAGAAATCCGAGCGGGCTCCGGAGCGGAGCGAG 117
QY      64 AGAGCGCGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 123
Db      118 GCGGCGCGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 177
QY      124 AGCGCACTGATTTGTCCTTGGGGCGGAGCGGAGCGGCGCGGAGATGAGCGCTGAT 183
Db      178 GGTGCACTGATTTGTCCTTGGGGCGGAGCGGAGCGGCGCGGAGATGAGCGCTGAT 237
QY      184 TAGCAAGTAAAGTAAAGAAACCAATGCTGATTTCCAAACCTTTGGTGGCAGCTG 243
Db      238 CAGCAAGTGAACCTAATAGAACCATGCTGATTTCCCAACCTTTGGTGGCAGCTG 297
QY      244 GATATCTGGGCGCATTAAGTGTAGAGAAAGAGAGAGATGATGATGATGATGATGAT 303
Db      298 GATGCTGGGCGCATTAAGTGTAGAGAAAGAGAGAGATGATGATGATGATGATGAT 357
QY      304 AAGCCAAATATCTGATTTGATTTGATGATGATGATGATGATGATGATGATGATGAT 363
Db      358 AAGCCGATAGCGGATTTTATCTGATGATGATGATGATGATGATGATGATGATGAT 417
QY      364 TTACCTCAACTGTTTATGACAGATATGAGCAGTATGATGATGATGATGATGATGAT 423
Db      418 TTACCTCAACTGTTTATGACAGATATGAGCAGTATGATGATGATGATGATGATGAT 477
QY      424 ATGATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTGAACTGAACT 483
Db      478 ATGATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTGAACTGAACT 537
QY      484 CAGCTACCCCTTGCACT 500
Db      538 CAGCTCCCTTCACACT 554

RESULT 15
BC062938      5385 bp      mRNA      linear      HTC 11-DEC-2003
LOCUS      MUS musculus intersectin (SH3 domain protein 1A), mRNA (cDNA clone
DEFINITION      IMAGE:6839463), containing frame-shift errors.
ACCESSION      BC062938
VERSION      BC062938.1 GI:38566052
KEYWORDS      HTC.
SOURCE      Mus musculus
ORGANISM      Mus musculus (house mouse)
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS      1 (bases 1 to 5385)
TITLE      Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
```

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marinina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stagleton, M., Soares, M.B., Bonaldo, M.F., Casavant, J.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carinini, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mallah, J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunatane, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.N., Sodergren, E.J., Lu, X., Gibbs, R.A., Fanhey, V., Helton, E., Kettlemen, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skelske, U., Smallus, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 5385)
Strausberg, R.
Direct Submission
Submitted (26-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-rc@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: M. Bento Soares, University of Iowa
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas U. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Sitr, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: Plate: Row: Column: 0
This clone has the following problem: frame shifted.

FEATURES
source
1. 5385
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGB:4839463"
/tissue_type="Brain"
/clone_id="NIF_BMAP_GH0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"

ORIGIN
Query Match 73.5%; Score 367.4; DB 11; Length 5385;
Best Local Similarity 85.5%; Pfd. No. 1e-84;
Matches 425; Conservative 0; Mismatches 61; Indels 11; Gaps 1;

QY 15 GTACGGCGGCTCGCGAGGAAGATCCCGAGCGGCTCCGGGACGGA-----CAG 63
DB 58 GTACGGCGGCTCGCGAGGAAGATCCCGAGCGGCTCCGGGACGGAAGGCGAG 117
QY 64 AGAGGGCGGCGGAGATGATGTGCGGGGCTCGGCTCCTGCTCCCGACGCGCGCTG 123
DB 118 GCGGGGCGGCGGAGATGATGTGCGGGGCTCGGAGCTCGGCTTCTCCGCGCGGCTGCG 177
QY 124 AGCGGACATGATTGTGCTCGGCGGCGGACCGCGGACCGCGGAGATGAGCGCTGAT 183
DB 178 GGCTGACTGATTGTGTAGAGGCGGCGCGCGGACCGCGCGGAGATGAGCGCTGAT 237

QY 184 TAGCAAGTAAAGTAACAGAACCAATGCGCTCACTTCCAAACCTTTGGTGGACGCTG 243
DB 238 CAGCAAGGTGAACGTAATGAACCATGCGCTCACTTCCAAACCTTTGGTGGACGCTG 297
QY 244 GATATCTGGGCGCATTAAGTATAGAGGAAGAGGACATGATCAGAGTTCATGTTTA 303
DB 298 GATGCTGGGCGATTAAGTATAGAGGAAGAGGACATGATCAGAGTTCATGTTTA 357
QY 304 AAGCCATATCTGATATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 363
DB 358 AAGCCGATGCGGAGATTTATCTGATCTGATCTGATCTGATCTGATCTGATCTG 417
QY 364 TTACCTCAACCTGTTTACACAGATATAGGCGACTGATGATGATGATGATGATG 423
DB 418 TTACCTCAACCTGTTTACACAGATATAGGCGACTGATGATGATGATGATGATG 477
QY 424 ATGATCAAGTGAAGTATTCATAGCTATGAACTTATCAAACTGAAGCTACAGGATAT 483
DB 478 ATGATCAAGTGAAGTATTCATAGCTATGAACTTATCAAACTGAAGCTACAGGATAT 537
QY 484 CAGCTACCTCTGCACT 500
DB 538 CAGCTCCTCCCACT 554

Search completed: July 1, 2004, 13:01:15
Job time : 2663 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2004, 08:47:35 ; Search time 388 Seconds
(without alignments)
5474.482 Million cell updates/sec

Title: US-09-720-934-1_COPY_1_500
Perfect score: 500
Sequence: 1 caaagaatccggtacgg.....taccgctaccctctgcact 500

Scoring table: IDENTITY NUC
Gapop 10% , Gapext 1.0

Searched: 3373863 segs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29JAN04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	5199	AA234570	Aa234570 Human SH3
2	486	97.2	877	AAK93179	AaK93179 Human CDN
3	486	97.2	877	AAK91610	AaK91610 Human CDN
4	486	97.2	1676	AAK94611	AaK94611 Human CDN
5	486	97.2	2131	AAK94139	AaK94139 Human full
6	486	97.2	2131	AAH16578	AaH16578 Human CDN
7	486	97.2	3466	AA163825	Aa163825 Human pol
8	486	97.2	5195	AA234572	Aa234572 Human SH3
9	486	97.2	5458	AA234571	Aa234571 Human SH3
10	484.4	96.9	7435	AA584763	Aa584763 DNA encod
11	477.4	94.3	676	AAH08146	AaH08146 Human CDN
12	444.4	88.9	3319	ABK43498	ABK43498 DNA encod
13	428	85.6	2079	AA234573	Aa234573 Human SH3
14	357	71.4	5082	AA239008	Aa239008 Mouse Ese
15	357	71.4	5738	AA239024	Aa239024 Mouse Ese
16	345	69.0	568	AA163919	Aa163919 Human pol
17	345	69.0	568	AA531621	Aa531621 CDNA encod
18	345	69.0	568	ABK43814	ABK43814 DNA encod
19	241.8	48.4	3723	AA239009	Aa239009 Mouse Ese
20	241.8	48.4	5144	AA239025	Aa239025 Mouse Ese
21	154.4	30.6	6103	AAK52332	AAK52332 Human pol
22	152.8	30.6	967	AA163832	Aa163832 Human pol
23	152.8	30.6	967	AA531616	AA531616 CDNA encod

24	152.8	30.6	967	4	ABK43821	ABK43821 DNA encod
25	152.8	30.6	5828	6	AA147247	AA147247 Allergic
26	140.4	28.1	4625	3	AA239010	Aa239010 Mouse Ese
27	140.4	28.1	6014	3	AA239026	Aa239026 Mouse Ese
28	137.8	27.6	531	3	AA569695	Aa569695 Human ova
29	137.8	27.6	531	6	ABN72589	ABn72589 Ovarian c
30	137.8	27.6	531	8	ADA08754	Ada08754 Human ova
31	137.8	27.6	2017	3	ABN72656	ABn72656 Ovarian c
32	137.8	27.6	2017	6	ABN72656	ABn72656 Human ova
33	137.8	27.6	2017	8	ADA08821	Ada08821 Human ova
34	133.6	26.7	480	3	AAA69750	AaA69750 Human ova
35	133.6	26.7	480	6	ABN72644	ABn72644 Ovarian c
36	133.6	26.7	480	8	ADA08809	Ada08809 Human ova
37	133.6	26.7	3592	3	AA239011	Aa239011 Mouse Ese
38	133.6	26.7	4975	3	AA239027	Aa239027 Mouse Ese
39	122.8	24.6	982	6	ABQ32431	ABq32431 Oligonuci
40	122.8	24.6	982	6	ABQ32430	ABq32430 Oligonuci
41	122.4	24.5	462	8	ACH34516	Ach34516 Human end
42	113.2	22.6	4447	4	AA502055	Aa502055 DNA encod
43	109	21.8	292	4	AA122800	Aa122800 Probe #12
44	109	21.8	292	4	ABA67887	AbA67887 Human foe
45	109	21.8	292	4	AA148103	AA148103 Probe #16

ALIGNMENTS

RESULT 1
ID AA234570 standard; CDNA; 5199 BP.
XX AC AA234570;
DT 01-FEB-2000 (first entry)
XX DE Human SH3D1A CDNA.
XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
XX megakaryocytic abnormality; myeloproliferative disorder;
XX platelet disorder; neural disorder; thrombocytopenia;
XX haematopoietic disorder; cognitive dysfunction; microcephaly;
XX lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX Homo sapiens.
XX OS
XX FH Key Location/Qualifiers
XX CDS 208..3642
XX FT /*tag= a
XX PN
XX MO9953062-A2.
XX PD 21-OCT-1999.
XX PF 16-APR-1999; 99WO-US008371.
XX PR 16-APR-1999; 98US-0082007P.
XX PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX PI Korenberg JR, Chen X;
XX DR MPI: 1999-633829/54.
XX PT P-PSDB; AA132154.
XX Nucleic acid from the human SH3D1A gene and its products, useful for the
XX diagnosis and treatment of myeloproliferative disorders and leukemia.
XX PS Claim 2; Fig 5; 99pp; English.
XX This is the nucleotide sequence of full-length CDNA corresponding to a
XX novel human SH3 gene, termed the SH3D1A gene, that contributes to the
XX development of platelets and the pathogenesis of leukemias, both in
XX general and in particular those involving the megakaryocytic lineage. The

SH3D1A gene maps to the small candidate region for low platelets on chromosome 21. Sequencing of 5 different sizes of cDNA clone (see AA234570-74) suggests that at least 3 isoforms exist. The invention provides methods for the diagnosis and treatment of megakaryocytic abnormalities, myeloproliferative disorder, platelet disorder, acute leukaemia, neural disorders, thrombocytopenia, platelet disorder on chromosome 21, low platelets in deletion for 21, association of gains in chromosome 21 with leukaemias, neural abnormalities, dysfunction of gains in disorders including brain malformations and corresponding cognitive dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are also provided for: suppressing cells unable to regulate themselves; screening for a somatic alteration in the SH3D1A gene; monitoring the progress and adequacy of a treatment; monitoring tumour risk progress or megakaryocytic abnormality, myeloproliferative disorder, haematopoietic disorder, platelet disorder or leukaemia; and treatment of a subject (including a prenatal subject) having megakaryocytic abnormality, myeloproliferative disorder, platelet disorder, leukaemia or neural disorder using a nucleic acid that expresses SH3D1A or its antisense nucleic acid

Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

Query Match 100.0%; Score 500; DB 2; Length 5199;

Best Local Similarity 100.0%; Pred. No. 3e-142; Mismatches 0; Gaps 0;

Matches 500; Conservative 0; Indels 0; Gaps 0;

1 CAAAGAGATTCGGGATACGGGCTCGCGAGAGAAATCCCGAGCGGGCTCCGGGACGGA 60

1 CAAAGAGATTCGGGATACGGGCTCGCGAGAGAAATCCCGAGCGGGCTCCGGGACGGA 60

61 CAGAGAGCGGGGCGGGGATGCTGCGGGGCTCGGGCTCCCTCCAGCGGCGC 120

61 CAGAGAGCGGGGCGGGGATGCTGCGGGGCTCGGGCTCCCTCCAGCGGCGC 120

122 GTGAGGCGGACGATTTGTCCTCGGGGCGGACGCGGACCCCGGAGATAGGCGTC 180

121 GTGAGGCGGACGATTTGTCCTCGGGGCGGACGCGGACCCCGGAGATAGGCGTC 180

181 GATTAGCAAGGTAAAGTAACAGAACCATGCTCAATTTCCAAACCTTTTGGTGGCAG 240

181 GATTAGCAAGGTAAAGTAACAGAACCATGCTCAATTTCCAAACCTTTTGGTGGCAG 240

241 CTGGATATCTGGGCGCACTGTAAGAGAAAGCAAGCATGATGAGAGTTCCATAGT 300

241 CTGGATATCTGGGCGCACTGTAAGAGAAAGCAAGCATGATGAGAGTTCCATAGT 300

301 TTAAGCCAAATCTGGATTCATTACTGGTGATCAAGCTAGAACTTTTTCATCT 360

301 TTAAGCCAAATCTGGATTCATTACTGGTGATCAAGCTAGAACTTTTTCATCT 360

361 GGGTTAAGCTCAACCTGTTTAAACAGATATGAGGCACTGATGATGATGATGGA 420

361 GGGTTAAGCTCAACCTGTTTAAACAGATATGAGGCACTGATGATGATGATGGA 420

421 ACAAATGATCAAGTGAAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGA 480

421 ACAAATGATCAAGTGAAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGA 480

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

481 TATCAGCTACCTCTGCACCT 500

Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

Homo sapiens.

EP1130094-A2.

05-SEP-2001.

07-JUL-2000; 2000EP-00114089.

08-JUL-1999; 99JP-00194486.

11-JAN-2000; 2000JP-00118774.

02-MAY-2000; 2000JP-00183765.

(HELI-) HELIX RES INST.

Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y, Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

WPI; 2001-524255/58.

830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation.

Example 11; SEQ ID NO 1639; 1380bp + Sequence Listing; English.

The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5' and 3' ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence was used as the representative sequence from a human clone which was used in homology searches to identify the clone. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO

Sequence 877 BP; 208 A; 228 C; 223 G; 213 T; 0 U; 5 Other;

Query Match 97.2%; Score 486; DB 4; Length 877;

Best Local Similarity 100.0%; Pred. No. 2.4e-138; Mismatches 486; Conservative 0; Indels 0; Gaps 0;

15 GTACGGCGGCTCGCGAGAGAAATCCCGAGCGGGCTCGGAGCGAGAGAGCGGGCG 74

12 GTACGGCGGCTCGCGAGAGAAATCCCGAGCGGGCTCGGAGCGAGAGAGCGGGCG 71

75 GGGATGATGATGCGGGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCG 134

72 GGGATGATGATGCGGGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCG 131

135 TTTGTCCTCGGGGCGGACGCGGACCCCGCGAGATAGGCGTGCATTTAGCAAGTAA 194

132 TTTGTCCTCGGGGCGGACGCGGACCCCGCGAGATAGGCGTGCATTTAGCAAGTAA 191

195 AAGTAACAGAACCAATGCTCAATTTCCAAACCTTTTGGTGGAGCTGATATCTGGCC 254

192 AAGTAACAGAACCAATGCTCAATTTCCAAACCTTTTGGTGGAGCTGATATCTGGCC 251

255 CATTAAGTGAAGAGAGAGAGAGAGATGATGAGAGCTTCAATGTTTAAAGCCAAATTC 314

252 CATTAAGTGAAGAGAGAGAGAGATGATGAGAGCTTCAATGTTTAAAGCCAAATTC 311

315 TGGATTCATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 374

312 TGGATTCATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 371

375 TGTTTTGAAGAGAGATGAGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 434

372 TGTTTTGAAGAGAGATGAGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 431

372 TGTTTTGAAGAGAGATGAGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 431

372 TGTTTTGAAGAGAGATGAGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 431

372 TGTTTTGAAGAGAGATGAGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 431

372 TGTTTTGAAGAGAGATGAGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 431

372 TGTTTTGAAGAGAGATGAGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 431

QY 435 GGAGTTTCCATAGCTATGAACTTATCAAACTGAGCTACAGATATCAGTACCCTC 494
DB 432 GGAGTTTCCATAGCTATGAACTTATCAAACTGAGCTACAGATATCAGTACCCTC 491
QY 495 TGCACT 500
DB 492 TGCACT 497

RESULT 3
AK91610
ID AK91610 standard; cDNA; 877 BP.
AC AK91610;
XX
XX 06-NOV-2001 (first entry)
DE Human cDNA 5'-end sequence, SEQ ID NO: 70.
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
XX EP1130094-A2.
XX 05-SEP-2001.
XX 07-JUL-2000; 2000EP-00114089.
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 2; SEQ ID NO 70; 1380bp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesized by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is the nucleotide sequence of the 5'-end of
CC a cDNA provided in the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in CD-
CC ROM format directly from EPO
XX
SQ Sequence 877 BP; 208 A; 228 C; 223 G; 213 T; 0 U; 5 Other;

Query Match 97.2%; Score 486; DB 4; Length 877;
Best Local Similarity 100.0%; Pred. No. 2,4e-138;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGGAGGAAGATCCGAGCGGCTCGGAGCGAGCAGAGAGGCGGCG 74
DB 12 GTACGGCGGCTCGGAGGAAGATCCGAGCGGCTCGGAGCGAGCAGAGAGGCGGCG 71
QY 75 GGAGTGTGTGCGGGGCTCGGCTCCTGCTCCCTCCAGCGCGGTGAGGGCGCACTGA 134
DB 72 GGAGTGTGTGCGGGGCTCGGCTCCTGCTCCCTCCAGCGCGGTGAGGGCGCACTGA 131

QY 135 TTTGTCCCTGGGGGCGGAGCGGAGCCGCCCGGAGATGAGGCGTGCATTAGCAAGGTAA 194
DB 132 TTTGTCCCTGGGGGCGGAGCGGAGCCGCCCGGAGATGAGGCGTGCATTAGCAAGGTAA 191
QY 195 AAGTACAGAAACCATGAGCTCAGTTTCCAAACCTTTTGGTGGAGCTGATATCTGGGC 254
DB 192 AAGTACAGAAACCATGAGCTCAGTTTCCAAACCTTTTGGTGGAGCTGATATCTGGGC 251
QY 255 CATTAAGTGTGAGGAAAGAGCGAAAGCATGATCAGCAGTTTCCATTGTTAAAGCCAAATTC 314
DB 252 CATTAAGTGTGAGGAAAGAGCGAAAGCATGATCAGCAGTTTCCATTGTTAAAGCCAAATTC 311
QY 315 TGGATTGATACGTCGTATGAACTTATCAAACTGAGCTACAGATATCAGTACCCTC 374
DB 312 TGGATTGATACGTCGTATGAACTTATCAAACTGAGCTACAGATATCAGTACCCTC 371
QY 375 TGGTTAGCAGATATGAGGAGCTAGCTGACATGATATGATGAGAAATGATCAAGT 434
DB 372 TGGTTAGCAGATATGAGGAGCTAGCTGACATGATATGATGAGAAATGATCAAGT 431
QY 435 GGAGTTTCCATAGCTATGAACTTATCAAACTGAGCTACAGATATCAGTACCCTC 494
DB 432 GGAGTTTCCATAGCTATGAACTTATCAAACTGAGCTACAGATATCAGTACCCTC 491
QY 495 TGCACT 500
DB 492 TGCACT 497

RESULT 4
AK94611
ID AK94611 standard; cDNA; 1676 BP.
AC AK94611;
XX
XX 07-NOV-2001 (first entry)
DE Human cDNA, SEQ ID NO: 3565.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX Homo sapiens.
XX EP1130094-A2.
XX 05-SEP-2001.
XX 07-JUL-2000; 2000EP-00114089.
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX P-PSDB; AAM93676.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Disclosure; SEQ ID NO 3565; 1380bp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesized by the oligo-capping method. The primers

AAH16578;
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:15658.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
XX
PS Claim 8; SEQ ID NO 15658; 2537bp + Sequence listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH16742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 2131 BP; 691 A; 448 C; 572 G; 420 T; 0 U; 0 Other;

Query Match 97.2%; Score 486; DB 4; Length 2131;
Best Local Similarity 100.0%; Pred. No. 3.8e-138;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GAGCGCGCTCGGAGGAAGATCCGAGCGGGCTCCGGAGCGAGACAGAGAGCGGGG 74
DB 154 GACGCGGCTCGGAGGAAGATCCGAGCGGGCTCCGGAGCGAGACAGAGAGCGGGG 213
QY 75 GGGATGATGTGCGGGGCTCGGCTCTGTGCTCCCTCCAGCGCGCGGTGAGCGGACTGA 134
DB 214 GGGATGATGTGCGGGGCTCGGCTCTGTGCTCCCTCCAGCGCGCGGTGAGCGGACTGA 273

QY 135 TTGTCCCTGGGGGCGGAGCGGAGCCCGCGGAGATGAGGGGTGATTAGCAAGTTA 194
DB 274 TTGTCCCTGGGGGCGGAGCGGAGCCCGCGGAGATGAGGGGTGATTAGCAAGTTA 333
QY 195 AAGTAAAGAACCATGGCTCAGTTTCCAGACCTTTTGTGGAGCCTGTGATATCTGGGC 254
DB 334 AAGTAAAGAACCATGGCTCAGTTTCCAGACCTTTTGTGGAGCCTGTGATATCTGGGC 393
QY 255 CATTAAGTGAAGAAAGAGCGGAGCATGATCAGCAGTTCCATGTTTAAAGCCATATC 314
DB 394 CATTAAGTGAAGAAAGAGCGGAGCATGATCAGCAGTTCCATGTTTAAAGCCATATC 453
QY 315 TGATTATATTCAGTGAATCAAGCTGAATCTTTTTCATCTGGGTACTCAACC 374
DB 454 TGATTATATTCAGTGAATCAAGCTGAATCTTTTTCATCTGGGTACTCAACC 513
QY 375 TGTTTAGCAGATATGAGCAGTACTGATGATGATGATGATGATGATGATGATGAT 434
DB 514 TGTTTAGCAGATATGAGCAGTACTGATGATGATGATGATGATGATGATGATGAT 573
QY 435 GGAAGTTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
DB 574 GGAAGTTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 633
QY 495 TGCACT 500
DB 634 TGCACT 639

RESULT 7
AA163825
ID AA163825 standard; cDNA; 3466 BP.
XX
XX AA163825;
DT 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 33.
DB
XX
KW Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; vnuclide;
KW fungicide; ophthalmological; cytostatic; immunosuppressive; neurotropic;
KW neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antitumor; antineoplastic; anticonvulsant; antibacterial;
KW antiparasitic; cardiant; gene therapy; cancer; immune disorder;
KW cardiovascular disorder; neurological disease; infection; human; ss.
XX
OS Homo sapiens.
XX
XX WO200155308-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001309.
PR
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 18-MAR-2000; 2000US-0190076P.
PR 17-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 23-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 08-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 12-SEP-2000; 2000US-0232081P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234423P.
 PR 21-SEP-2000; 2000US-0234474P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236602P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237340P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 PI WPI: 2001-488781/53.
 DR P-PSDB; AAM43519.
 XX
 PT New isolated nucleic acids and polypeptides, useful for diagnosing,
 PT treating and/or preventing human diseases and disorders.
 XX
 PS Claim 1. SEQ ID NO 33; 664pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AA163803-AA164012) and
 CC the encoded proteins (AAM43497-AAM4360) useful for preventing, treating
 CC or ameliorating medical conditions e.g. by protein or gene therapy. The
 CC genes were isolated from a range of human tissues disclosed in the
 CC specification. The nucleic acids, proteins, antibodies and (ant)agonists
 CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
 CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,
 CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
 CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,
 CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
 CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections. Note: The sequence data for this patent did not

CC form part of the printed specification, but was obtained in electronic
 CC format directly from MPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3466 BP; 1056 A; 829 C; 861 G; 710 T; 0 U; 10 Other;

Query Match 97.2%; Score 486; DB 4; Length 3466;

Best Local Similarity 100.0%; Pred. No. 4.8e-138; Indels 0; Gaps 0;

Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGGCTCCGGAGCGAGAGAGCGGGCG 74
 DB 84 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGGCTCCGGAGCGAGAGAGCGGGCG 143
 QY 75 GGGATGCTGTCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCTGAGCGGACTGA 134
 DB 144 GGGATGCTGTCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCTGAGCGGACTGA 203
 QY 135 TTTTCCCTCGGGGGCGGAGCGCGGAGCCCGCGAGATGAGAGCGGCTGAGCAAGGTAA 194
 DB 204 TTTTCCCTCGGGGGCGGAGCGCGGAGCCCGCGAGATGAGAGCGGCTGAGCAAGGTAA 263
 QY 195 AAGTAACAGAACCATGCTCAGATTCCAAACCTTTTGGTGGCAGCCCTGATATCTGGGC 254
 DB 264 AAGTAACAGAACCATGCTCAGATTCCAAACCTTTTGGTGGCAGCCCTGATATCTGGGC 323
 QY 255 CATTAAGTATGAGAGAGAGAGAGAGAGATGATGAGATTCATATGTTAAAGCCCAATTC 314
 DB 324 CATTAAGTATGAGAGAGAGAGAGAGAGATGATGAGATTCATATGTTAAAGCCCAATTC 383
 QY 315 TGGATTCATTAATCTGGTATCAAGCTAGAGAACTTTTTCATCTGGGTTACCTCAAC 374
 DB 384 TGGATTCATTAATCTGGTATCAAGCTAGAGAACTTTTTCATCTGGGTTACCTCAAC 443
 QY 375 TGTTTTACAGACAGATATGGGCACTAGCTGACATGATATATGATGAGAGATGATCAAGT 434
 DB 444 TGTTTTACAGACAGATATGGGCACTAGCTGACATGATATATGATGAGAGATGATCAAGT 503
 QY 435 GGAGTTTTCATGCTATGAAACTTATCAAACTGAAAGCTACAGAGATTCAGCTACCTC 494
 DB 504 GGAGTTTTCATGCTATGAAACTTATCAAACTGAAAGCTACAGAGATTCAGCTACCTC 563
 QY 495 TGGACT 500
 DB 564 TGGACT 569

RESULT 8
 AA234572 standard: cDNA, 5195 BP.
 ID AA234572
 XX
 AC AA234572;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 11.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 239..3886
 ET /*tag= a
 XX
 PN WO9953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WO-US008371.

XX
 PR 16-APR-1999; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SIGNAL HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI, 1999-633829/54.
 DR P-PSDB, AAY32156.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 10; 99pp; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukaemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;

Query Match 97.2%; Score 486; DB 2; Length 5195;
 Best Local Similarity 100.0%; Pred. No. 5.9e-138;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGGCTCCGGAGCGAGAGAGCGGGCG 74
 DB 46 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGGCTCCGGAGCGAGAGAGCGGGCG 105
 QY 75 GGGATGCTGTCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGGCTGAGCGGACTGA 134
 DB 106 GGGATGCTGTCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGGCTGAGCGGACTGA 165
 QY 135 TTTTCCCTCGGGGGCGGAGCGCGGAGCCCGCGAGATGAGAGCGTCAATTAGCAAGGTAA 194
 DB 166 TTTTCCCTCGGGGGCGGAGCGCGGAGCCCGCGAGATGAGAGCGTCAATTAGCAAGGTAA 225
 QY 195 AAGTAACAGAACCATGCTCAGATTCCAAACCTTTTGGTGGCAGCCCTGATATCTGGGC 254
 DB 226 AAGTAACAGAACCATGCTCAGATTCCAAACCTTTTGGTGGCAGCCCTGATATCTGGGC 285
 QY 255 CATTAAGTATGAGAGAGAGAGAGAGAGATGATGAGATTCATATGTTAAAGCCCAATTC 314
 DB 286 CATTAAGTATGAGAGAGAGAGAGAGAGATGATGAGATTCATATGTTAAAGCCCAATTC 345
 QY 315 TGGATTCATTAATCTGGTATCAAGCTAGAGAACTTTTTCATCTGGGTTACCTCAAC 374
 DB 346 TGGATTCATTAATCTGGTATCAAGCTAGAGAACTTTTTCATCTGGGTTACCTCAAC 405
 QY 375 TGTTTTACAGACAGATATGGGCACTAGCTGACATGATATATGATGAGAGATGATCAAGT 434
 DB 406 TGTTTTACAGACAGATATGGGCACTAGCTGACATGATATATGATGAGAGATGATCAAGT 465

QY 435 GGAGTTTTCATAGCTATGAACCTTATCAACTGAGCTACAAGATATCAGTACCCCTC 494
 DB 466 GGAGTTTTCATAGCTATGAACCTTATCAACTGAGCTACAAGATATCAGTACCCCTC 525
 QY 495 TGCACCT 500
 DB 526 TGCACCT 531

RESULT 9
 AA234571
 ID AA234571 standard; cDNA; 5458 BP.
 XX
 AC AA234571;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 21.
 XX
 SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 KM megakaryocytic abnormality; myeloproliferative disorder;
 KM platelet disorder; neural disorder; thrombocytopenia;
 KM haemopoietic disorder; cognitive dysfunction; microcephaly;
 KM lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 267..3929
 FT /*tag= a
 XX
 PN W09953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WO-US008371.
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI; 1999-633829/54.
 DR P-PSDB; AAY32155.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 8; 99P; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 21)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukemia, neural disorders, thrombocytopenia,
 CC association of gains in chromosome 21, low platelets in deletion for 21,
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukemia or neural disorder using a nucleic acid that

CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;
 Query Match 97.2%; Score 486; DB 2; Length 5458;
 Best Local Similarity 100.0%; Pred. No. 6,1e-138;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTGCGAGAGAGAAATCCCGAGCGGCTCCGGAGCGAGACAGAGAGCGGGCG 74
 DB 74 GTACGGCGGCTGCGAGAGAGAAATCCCGAGCGGCTCCGGAGCGAGACAGAGAGCGGGCG 133
 QY 75 GGGATGGTGGCGGGGCTGGCGCTCCCTGCGCTCCCGACCGGCGCTGAGCGGCACTGA 134
 DB 134 GGGATGGTGGCGGGGCTGGCGCTCCCTGCGCTCCCGACCGGCGCTGAGCGGCACTGA 193
 QY 135 TTTGTCCTGGGCGGCGAGCGCGGACCGCGCGAGATGAGCGTCGATTAGCAAGTAA 194
 DB 194 TTTGTCCTGGGCGGCGAGCGCGGACCGCGCGAGATGAGCGTCGATTAGCAAGTAA 253
 QY 195 AAGTAAACAGAACATGCGCTGAGTTTCCACACACTTTTGTGGGAGCTGAGATCTGGGC 254
 DB 254 AAGTAAACAGAACATGCGCTGAGTTTCCACACACTTTTGTGGGAGCTGAGATCTGGGC 313
 QY 255 CATACCTGTAGAGAAAGAGCGAAGCATGATGACAGTTCCATTAGTTTAAAGCAATATC 314
 DB 314 CATACCTGTAGAGAAAGAGCGAAGCATGATGACAGTTCCATTAGTTTAAAGCAATATC 373
 QY 315 TGGATTCAATTACTGTGATCACTAGCAAACTTTTTCATCTGGTTACCTCAAC 374
 DB 374 TGGATTCAATTACTGTGATCACTAGCAAACTTTTTCATCTGGTTACCTCAAC 433
 QY 375 TGTTTTGAACAGATATGGGCACTAGCTGACATGAATATGATGAAAGATGATCAAGT 434
 DB 434 TGTTTTGAACAGATATGGGCACTAGCTGACATGAATATGATGAAAGATGATCAAGT 493
 QY 435 GGAGTTTTCATAGCTATGAACCTTATCAACTGAGCTACAAGATATCAGTACCCCTC 494
 DB 494 GGAGTTTTCATAGCTATGAACCTTATCAACTGAGCTACAAGATATCAGTACCCCTC 553
 QY 495 TGCACCT 500
 DB 554 TGCACCT 559

RESULT 10
 AAS84763
 ID AAS84763 standard; cDNA; 7435 BP.
 XX
 AC AAS84763;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #20567.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-6339362/73.
DR P-PSDB; ABG20576.

DR P-PSDB; ABG20576.

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

PS Claim 1; SEQ ID NO 20567; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. A5654197-A5934564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/publ/pct_sequences

Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;

Query Match 96.9%; Score 484.4; DB 5; Length 7435;

Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

15 GTACGGCGCTCGCGAGGAATCCCGACGGCTCCGGGACGGACAGAGAGCGCGG 74

QY	15	GTACGGCGGCTCGCAGGAAATATCCAGCGGGCTCCGGGACGACAGAGGGCGGGCG	74
Db	70	GTACGGCGGCTCGCAGGAAATATCCAGCGGGCTCCGGGACGACAGAGGGCGGGCG	129
QY	75	GGGATGCTGTGCGGGGCTGCGGCTCTGCTGCTCCCTCCAGCGGCGCTGAGGGCACTGCA	134
Db	130	GGGATGCTGTGCGGGGCTGCGGCTCTGCTGCTCCCTCCAGCGGCGCTGAGGGCACTGCA	189
QY	135	TTTGTCCCTGGGGCGCGCAGCGGACCGCGCCGGAGATGAGGGCTGCATTTAGCAAGTGA	194
Db	190	TTTGTCCCTGGGGCGCGCAGCGGACCGCGCCGGAGATGAGGGCTGCATTTAGCAAGTGA	249
QY	195	AAGTAAACGAACCATGGCTCAGTTTCCAAACACTTTTGGTGGAGCTGGATATCTGGGC	254
Db	250	AAGTAAACGAACCATGGCTCAGTTTCCAAACACTTTTGGTGGAGCTGGATATCTGGGC	309
QY	255	CATACTGTAGAGGAAAGCGGACGATGATCAGCAGTCCATGATTAAACCAATATC	314
Db	310	CATACTGTAGAGGAAAGCGGACGATGATCAGCAGTCCATGATTAAACCAATATC	369
QY	315	TGATATCTACTGCTGATCAAGCTAGAAACTTTTTTTTCAATCTGGGTAACTTCAACC	374
Db	370	TGATATCTACTGCTGATCAAGCTAGAAACTTTTTTTTCAATCTGGGTAACTTCAACC	429
QY	375	TGTTTACACAGATATGGGCACTGACGACATGAAATAATGATGAAGATGATCAACT	434
Db	430	TGTTTACACAGATATGGGCACTGACGACATGAAATAATGATGAAGATGATCAACT	489
QY	435	GGAGTTTTCATAGCTATGAAACTTATCAAACTGAAGCTACAAAGATATCACTACCTC	494
Db	490	GGAGTTTTCATAGCTATGAAACTTATCAAACTGAAGCTACAAAGATATCACTACCTC	549
QY	495	TGCACT 500	
Db	550	TGCACT 555	

RESULT 11

ID AAH08146 standard; cDNA; 676 BP.

AC AAH08146;

DT 26-JUN-2001 (first entry)

DE	Human cDNA clone (5'-primer)	SEQ ID NO:4981.

KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

05 Homo sapiens.

PN EP1074617-A2

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-00116126.

PR 29-JUL-1999; 99JP-00248036.

PR 11-JAN-2000; 2000JP-00118776.

PR 09-JUN-2000; 2000JP-00241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nish

XX

XX

PT	length	CDNAs defined
1	100	1
2	100	1
3	100	1
4	100	1
5	100	1
6	100	1
7	100	1
8	100	1
9	100	1
10	100	1
11	100	1
12	100	1
13	100	1
14	100	1
15	100	1
16	100	1
17	100	1
18	100	1
19	100	1
20	100	1
21	100	1
22	100	1
23	100	1
24	100	1
25	100	1
26	100	1
27	100	1
28	100	1
29	100	1
30	100	1
31	100	1
32	100	1
33	100	1
34	100	1
35	100	1
36	100	1
37	100	1
38	100	1
39	100	1
40	100	1
41	100	1
42	100	1
43	100	1
44	100	1
45	100	1
46	100	1
47	100	1
48	100	1
49	100	1
50	100	1
51	100	1
52	100	1
53	100	1
54	100	1
55	100	1
56	100	1
57	100	1
58	100	1
59	100	1
60	100	1
61	100	1
62	100	1
63	100	1
64	100	1
65	100	1
66	100	1
67	100	1
68	100	1
69	100	1
70	100	1
71	100	1
72	100	1
73	100	1
74	100	1
75	100	1
76	100	1
77	100	1
78	100	1
79	100	1
80	100	1
81	100	1
82	100	1
83	100	1
84	100	1
85	100	1
86	100	1
87	100	1
88	100	1
89	100	1
90	100	1
91	100	1
92	100	1
93	100	1
94	100	1
95	100	1
96	100	1
97	100	1
98	100	1
99	100	1
100	100	1

PT cDNAs.

PS Claim 1; SEQ ID NO 4981; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH01662 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95889 represent human amino acid sequences; and AAH11629 to AAH11632 represent polynucleotides, all of which are used in the exemplification of the present invention.

Sequence 676 BP; 167 A; 141 C; 228 G; 137 T; 0 U; 3 Other;

Query Match	94.38;	Score 471.4;	DB 4;	Length 676;
-------------	--------	--------------	-------	-------------

Matches 483; Conservative 0; Mismatches 2; Indels 1; Gaps 1.

15 GTACGGCGGCTCCGCGAGGAAGATCCCGAGCGGGCTCCGGGACGGACAGAGAGGCGGCG /4

QY 15 GTACGCGCGCTCCGCGAGGAAGATCCCGAGCGGGCTCCGGGACGGAACAGAGAGGCGGGCG 74
 |||||

DB 154 GTACGGGCGCTCCGACGAGAAATCCCGACGGGCTCCGGGACGAGACAGAGCGGGCG 213
QY 75 GGGATGGTGTGGGGGCTGCGCTCTGGCTCCAGCGGGCGGTGAGCGGCACACTGA 134
DB 214 GGGATGGTGTGGGGGCTGCGCTCTGGCTCCAGCGGGCGGTGAGCGGCACACTGA 273
QY 135 TTGTCCCTGGGGCGGCGCGCGGACCCCGCGAGATGAGGCGCTGATTAGCAAGTAA 194
DB 274 TTGTCCCTGGGGCGGCGCGCGGACCCCGCGAGATGAGGCGGTGATTAGCAAGTAA 333
QY 195 AAGTAACAGAACCATGGCTACTGTTTCCAAACCTTTTGGGCGACCTGGATATCTGGGC 254
DB 334 AAGTAACAGAACCATGGCTACTGTTTCCAAACCTTTTGGGCGACCTGGATATCTGGGC 393
QY 255 CATTAAGTGTAGAGAAAGAGAGAGATGATCAGAGATGATCAGAGATGATCAGAGATGATC 314
DB 394 CATTAAGTGTAGAGAAAGAGAGAGATGATCAGAGATGATCAGAGATGATCAGAGATGATC 453
QY 315 TGGATTCATTACTGGTATCAGAGATGATCAGAGATGATCAGAGATGATCAGAGATGATC 374
DB 454 TGGATTCATTACTGGTATCAGAGATGATCAGAGATGATCAGAGATGATCAGAGATGATC 513
QY 375 TGTTTAGCACACATATGGGCTAGTGTGATGATGATGATGATGATGATGATGATGATGAT 434
DB 514 TGTTTAGCACACATATGGGCTAGTGTGATGATGATGATGATGATGATGATGATGATGAT 572
QY 435 GGAGTTTTCATAGCTATGAACTTATCAAGTGAAGCTACAGATGATGATGATGATGATGAT 494
DB 573 TGAATTTTCATAGCTATGAACTTATCAAGTGAAGCTACAGATGATGATGATGATGATGAT 632
QY 495 TGCACCT 500
DB 633 TGCACCT 638
RESULT 12
ABK43498
ID ABK43498 standard; cDNA; 3319 bp.
XX
AC ABK43498;
XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding novel central nervous system protein #78.
XX
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiodysplasia;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200155318-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US001332.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 18-AUG-2000; 2000US-0226281P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 06-SEP-2000; 2000US-0230439P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234978P.
PR 25-SEP-2000; 2000US-0234980P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.

XX	Key	Location/Qualifiers
PH	CDS	136..2079
FT		/*tag= a
XX		
PN	MO9953062-A2.	
XX		
PD	21-OCT-1999.	
XX		
PF	16-APR-1999; 99WO-US008371.	
XX		
PR	16-APR-1998; 98US-0082007P.	
XX		
PA	(CEDA-) CEDARS SINAI HEALTH SYSTEM.	
XX		
PI	Korenberg JR, Chen X;	
XX		
DR	WPI, 1999-633829/54.	
XX		
PT	P-PsDB; AAY32157.	
XX		
PT	Nucleic acid from the human SH3D1A gene and its products, useful for the	
XX	diagnosis and treatment of myeloproliferative disorders and leukemia.	
XX	Claim 2; Fig 12; 99pp; English.	
XX		
XX	This is the nucleotide sequence of a cDNA clone, termed clone 5,	
CC	corresponding to a novel human SH3 gene, termed the SH3D1A gene, that	
CC	contributes to the development of platelets and the pathogenesis of	
CC	leukemias, both in general and in particular those involving the	
CC	megakaryocytic lineage. The SH3D1A gene maps to the small candidate	
CC	region for low platelets on chromosome 21. Sequencing of 5 different	
CC	sizes of cDNA clones from foetal brain (see AAZ34570-74) suggests that at	
CC	least 3 isoforms exist. The invention provides methods for the diagnosis	
CC	and treatment of megakaryocytic abnormality, myeloproliferative disorder,	
CC	platelet disorder, acute leukemia, neutral disorders, thrombocytopenia,	
CC	platelet disorder on chromosome 21, low platelets in deletion for 21,	
CC	association of gains in chromosome 21 with leukemias, neutral	
CC	abnormalities, dysfunctions and disorders including brain malformations	
CC	and corresponding cognitive dysfunctions, microcephaly, lissencephaly,	
CC	and colpocephaly. Methods are also provided for: suppressing cells unable	
CC	to regulate themselves; screening for a somatic alteration in the SH3D1A	
CC	gene; monitoring the progress and adequacy of a treatment; monitoring	
CC	tumour risk progress or megakaryocytic abnormality, myeloproliferative	
CC	disorder, haematopoietic disorder, platelet disorder or leukemia; and	
CC	treatment of a subject (including a prenatal subject) having	
CC	megakaryocytic abnormality, myeloproliferative disorder, platelet	
CC	disorder, leukemia or neutral disorder using a nucleic acid that	
CC	expresses SH3D1A or its antisense nucleic acid	
XX		
XX	Sequence 2079 BP; 700 A; 451 C; 495 G; 433 T; 0 U; 0 Other;	
SQ		
XX		
XX	Query Match 85.6%; Spore 428; DB 2; Length 2079;	
XX	Best Local Similarity 100.0%; Pred. No. 2.3e-120;	
XX	Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	73 CCGGAGATGATGTCGGGGGCTGCGCTCTCCGCTCCCGAGCGCGCTGAGCGCACT 132	
DB	1 CCGGAGATGATGTCGGGGGCTGCGCTCTCCGCTCCCGAGCGCGCTGAGCGCACT 60	
QY	133 GATTTGTCTCTGGGGGGGCGAGCCCGGAGATGAGGGCTGCATTGACAGGT 192	
DB	61 GATTTGTCTCTGGGGGGGCGAGCCCGGAGATGAGGGCTGCATTGACAGGT 120	
QY	193 AAAAGTACAGAACCTGCTGAGTTTCCACACCTTTTGGTGGGAGCGCTGATATCTGG 252	
DB	121 AAAAGTACAGAACCTGCTGAGTTTCCACACCTTTTGGTGGGAGCGCTGATATCTGG 180	
QY	253 GCCATTAATCTGAGAGAAAGAGCGAAGCATGATGATGACAGTTCATTGTTAAAGCCATA 312	
DB	181 GCCATTAATCTGAGAGAAAGAGCGAAGCATGATGATGACAGTTCATTGTTAAAGCCATA 240	
QY	313 TCTGATTCATTAATCTGATTAAGTAAAGCTTTTTCATCTGGGTTAACTCTCAA 372	
XX		

DB	241 TCTGATTCATTAATCTGATTAAGTAAAGCTTTTTCATCTGGGTTAACTCTCAA 300	
QY	373 CTTGTTTACCAAGATATGGGCACTAGCTGACATGATATATGAGAAATGATCAA 432	
DB	301 CTTGTTTACCAAGATATGGGCACTAGCTGACATGATATATGAGAAATGATCAA 360	
QY	433 GTGAGTTTCCATGAGTATGAACTTATCAACTGATGACAGGATTCAGCTACCC 492	
DB	361 GTGAGTTTCCATGAGTATGAACTTATCAACTGATGACAGGATTCAGCTACCC 420	
QY	493 TCTGCACT 500	
DB	421 TCTGCACT 428	
XX		
XX	RESULT 14	
XX	AAZ39008	
XX	AAZ39008 standard; cDNA; 5082 BP.	
XX		
AC	AAZ39008;	
XX		
DT	28-FEB-2000 (first entry)	
XX		
DE	Mouse Ees1 full length cDNA sequence.	
XX		
KW	Mouse; murine; Ees1; Ees2; endocytosis; vesicular trafficking;	
KW	regulation; actin cytoskeleton; detection; cancer; infection;	
KW	EH-domain and SH3-domain regulator of endocytosis; anticancer;	
XX	antiproliferative; antiviral; ss.	
XX		
OS	Mus sp.	
XX		
PN	WO9955728-A2.	
XX		
PD	04-NOV-1999.	
XX		
PF	27-APR-1999; 99WO-CA000375.	
XX		
PR	27-APR-1998; 98CA-02230201.	
PR	05-FEB-1999; 99US-0118739P.	
XX		
PA	(HSCR-) HSC RES & DEV LP.	
XX		
PI	Egan SE, Wang W, Sengar A;	
XX		
DR	WPI, 2000-052892/04.	
XX	P-PsDB; AAY57444.	
XX		
PT	New nucleic acid encoding Ees1 and 2 proteins, involved in regulation of	
PT	endocytosis, used e.g. for treating cancer or preventing viral infection.	
XX		
PS	Claim 6; Page 38-40; 99pp; English.	
XX		
XX	The present sequence encodes mouse Ees1. The present invention	
CC	specifically describes mammalian Ees1 and 2 proteins (I) and their splice	
CC	variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)	
CC	are involved in regulation of clathrin-mediated endocytosis (as a complex	
CC	with Ees1 protein), vesicular trafficking and actin cytoskeleton.	
CC	Generally (I) (or its (ant)agonists, mimetics, fragments and inactive	
CC	mutants); (II)-specific antibodies (Ab); sequences antisense to the (I)	
CC	polynucleotide; agents that downregulate expression of Ese genes or	
CC	antagonists of an Ese binding partner are used to treat diseases	
CC	associated with undesirable endocytosis and resulting changes in cellular	
CC	function. Particularly overexpression of Ees1 is used to block clathrin-	
CC	mediated endocytosis in vivo or in cell cultures, while administration of	
CC	(I) is used to promote endocytosis of selected cells. (Ant)agonists of	
CC	(I) or Ab are used to suppress abnormal proliferation of cells that can	
CC	be stimulated to proliferate by a growth factor receptor; and similar	
CC	compounds (also inactive Ese mutants) can be used to prevent viral	
CC	infection. Endocytosis may also be regulated, in vivo or in cell	
CC	cultures, by forming an Ese-Ees1 complex, then binding dynamin to the	
CC	complex. Generally conditions that can be treated include cancer;	
CC	abnormal cell division or migration; viral infection; or abnormal	

receptor signalling, tissue development or synaptic transmission

Sequence 5082 BP, 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;
Query Match 71.4%; Score 357; DB 3; Length 5082;
Best Local Similarity 85.5%; Pred. No. 2.1e-98;
Matches 425; Conservative 0; Mismatches 60; Indels 12; Gaps 2;

```
QY 15 GTACGGCGGCTCGCGAGGAAGAATCCGAGCGGGCTCCGGAGCGA-----CAG 63
DB 56 GTACGGCGGCTCGCGAGGAGGACATCCGAGCGGGCTCCGGAGCGGAGAGGAGAG 115
QY 64 AGAGCGCGCGGGGATGTGTGCGGGGCTGCGGCTCCCTCCGACGAGGCGCTG 123
DB 116 GCGGCGCGCGGGAGTGTGTGCGGGCTGCGAGCTCGGCTTCTCGC-GCGGCGTGC 174
QY 124 AGCGGCACTGATTTGTCCCTGGGCGGCGAGCGGACCGCGCGAGATGAGGCGT 183
DB 175 GCGTGCATGATTTGTGTGAGGGGGCGCGCGCGACCGCGGAGATGAGGCGT 234
QY 184 TAGCAAGGTAAAGTACAGAACCATGCTCAGTTTCCAAACCTTTGGTGGAGCTG 243
DB 235 CAGCAAGGTGAACGTAATAGAACCATGCTCAGTTTCCCAACCTTTGGTGGAGCTG 294
QY 244 GATATCTGGGCGCATTAAGTAGAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 303
DB 295 GATGCTGGGCGCATTAAGTAGAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 354
QY 304 AAGCCATATCTGATTCATTAAGTATGATCAAGGTAGAACTTTTTCATTCATG 363
DB 355 AAGCGATAGCGGGATTTATTAAGTATGATCAAGGTAGAACTTTTTCATTCATG 414
QY 364 TTACCTCACTGTTTGTGCAAGATATGAGGAGCTAGTGAAGTATATGATGAGA 423
DB 415 TTACCTCACTGTTTGTGCAAGATATGAGGAGCTAGTGAAGTATATGATGAGA 474
QY 424 ATGATCAAGTGAAGTTTTCATTAAGTATGAGGAGCTAGTGAAGTATATGATGAGA 483
DB 475 ATGATCAAGTGAAGTTTTCATTAAGTATGAGGAGCTAGTGAAGTATATGATGAGA 534
QY 484 CAGCTACCTCTGCACT 500
DB 535 CAGCTACCTCTGCACT 551
```

RESULT 15

AAZ39024
ID AAZ39024 standard; cDNA; 5738 BP.

AC AAZ39024;

DT 28-FEB-2000 (first entry)

DE Mouse Esell cDNA sequence.

KM Mouse; murine; Esel; Ees2; endocytosis; vesicular trafficking;

KW regulation; actin cytoskeleton; detection; cancer; infection;

KW EH-domain and SH3-domain regulator of endocytosis; anticancer;

KW antiproliferative; antiviral; ss.

OS Mus sp.

PN W09955728-A2.

PD 04-NOV-1999.

PF 27-APR-1999; 99WO-CA000375.

PR 27-APR-1998; 98CA-02230201.

PR 05-FEB-1999; 99US-0118739P.

PA (HSCR-) HSC RES & DEV LP.

PI Egan SE, Wang W, Sengar A;
XX WP1; 2000-052802/04.
DR P-PSDB; AAY57449.

PT New nucleic acid encoding Esel and 2 proteins, involved in regulation of
endocytosis, used e.g. for treating cancer or preventing viral infection.
XX Claim 6; Page 56-59; 99p; English.

The present invention specifically describes mammalian Esel and 2
proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
regulator of endocytosis). (I) are involved in regulation of clathrin-
mediated endocytosis (as a complex with Espl protein), vesicular
trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
sequences antisense to the (I) polynucleotide; agents that downregulate
expression of Ese genes or antagonists of an Ese binding partner are used
to treat diseases associated with undesirable endocytosis and resulting
changes in cellular function. Particularly overexpression of Esel is used
to block clathrin-mediated endocytosis in vivo or in cell cultures, while
administration of (I) is used to promote endocytosis of selected cells.
(Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
cells that can be stimulated to proliferate by a growth factor receptor;
and similar compounds (also inactive Ese mutants) can be used to prevent
viral infection. Endocytosis may also be regulated, in vivo or in cell
cultures, by forming an Ese-Espl complex, then binding dynamitin to the
complex. Generally conditions that can be treated include cancer;
abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission. The
present sequence represents mouse Esel cDNA sequence

Sequence 5738 BP, 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;

Query Match 71.4%; Score 357; DB 3; Length 5738;
Best Local Similarity 85.5%; Pred. No. 2.2e-98;
Matches 425; Conservative 0; Mismatches 60; Indels 12; Gaps 2;

```
QY 15 GTACGGCGGCTCGCGAGGAAGAATCCGAGCGGGCTCCGGAGCGA-----CAG 63
DB 57 GTACGGCGGCTCGCGAGGAGGACATCCGAGCGGGCTCCGGAGCGGAGGAGGAG 116
QY 64 AGAGCGCGCGGGGATGTGTGCGGGGCTGCGGCTCCCTCCGAGCGGCGGCTG 123
DB 117 GCGGCGCGCGGGAGTGTGTGCGGGGCTGCGAGCTCGGCTTCTCGC-GCGGCGTGC 175
QY 124 AGCGGCACTGATTTGTCCCTGGGCGGCGAGCGGACCGCGCGAGATGAGGCGT 183
DB 176 GCGTGCATGATTTGTGTGAGGGGGCGCGCGCGACCGCGCGGAGATGAGGCGT 235
QY 184 TAGCAAGGTAAAGTACAGAACCATGCTCAGTTTCCAAACCTTTGGTGGAGCTG 243
DB 236 CAGCAAGGTGAACGTAATAGAACCATGCTCAGTTTCCCAACCTTTGGTGGAGCTG 295
QY 244 GATATCTGGGCGCATTAAGTAGAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 303
DB 296 GATGCTGGGCGCATTAAGTAGAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 355
QY 304 AAGCCATATCTGATTCATTAAGTATGATCAAGGTAGAACTTTTTCATTCATG 363
DB 356 AAGCGATAGCGGGATTTATTAAGTATGATCAAGGTAGAACTTTTTCATTCATG 415
QY 364 TTACCTCACTGTTTGTGCAAGATATGAGGAGCTAGTGAAGTATATGATGAGA 423
DB 416 TTACCTCACTGTTTGTGCAAGATATGAGGAGCTAGTGAAGTATATGATGAGA 475
QY 424 ATGATCAAGTGAAGTTTTCATTAAGTATGAGGAGCTAGTGAAGTATATGATGAGA 483
DB 476 ATGATCAAGTGAAGTTTTCATTAAGTATGAGGAGCTAGTGAAGTATATGATGAGA 535
QY 484 CAGCTACCTCTGCACT 500
DB 536 CAGCTACCTCTGCACT 552
```

Fri Jul 2 10:23:33 2004

Search completed: July 1, 2004, 11:40:53
Job time : 393 secs

us-09-720-934-1_copy_1_500.rng

Page 14

RESULT 2
US-09-338-933-5
; Sequence 5, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lyn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-5

Query Match 27.6%; Score 137.8; DB 4; Length 531;
Best Local Similarity 64.7%; Pred. No. 1.1e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 175 GCGGTGATTAGCAAGTAAAGTAAACAGAACCATGGCTCACTTTCCAAACCTTTTGGT 234
DB 9 GCGTGAGAGCTGCAAGAAAGATCAGAGATCATGATGGCTCAGTTTCCACAGCATGAT 68
QY 235 GGCAGCCTGGATATCTGGGCGCATTAAGAGAAAGCGGAAGCATGATCAGCATGTC 294
DB 69 GGAAGGCGCAAAATATGAGGCTATTAATCATCTGAAAGCATTAAGCATGATTAAGCATGTT 128
QY 295 CATAGTTTAAAGCAATATCTGATTCATTAAGTATGATCAAGTAACTTTTCTTTT 354
DB 129 GATTAACCTCAAACTTCAAGAGGTTACATACAGGATCAGCCGTAATTTTCTTA 188
QY 355 CAATCTGGGTAACTCTCACTTTTGAACAGATATGGGCACTAGCTGACATGAATAAT 414
DB 189 CAGTCAGGTCGCGGCGCGGCTTTTGAATATGGGCTTATCAGATCTGAACAG 248
QY 415 GATGGAAGATGATCAAGTGAAGTTTTCATAGCTATGAACTTTCAAACTGAAGCTA 474
DB 249 GATGGGAAGATGAGACCAAGATCTCTATAGCTATGAACTTTCATCAAGTTAAAGTTG 308
QY 475 CAAGATATCAGCTACC 491
DB 309 CAGGGCCCAACAGCTGCC 325

RESULT 3
US-09-215-681-5
; Sequence 5, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF OVARIAN CANCER
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-5

Query Match 27.6%; Score 137.8; DB 4; Length 531;
Best Local Similarity 64.7%; Pred. No. 1.1e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 175 GCGGTGATTAGCAAGTAAAGTAAACAGAACCATGGCTCACTTTCCAAACCTTTTGGT 234
DB 9 GCGTGAGAGCTGCAAGAAAGATCAGAGATCATGATGGCTCAGTTTCCACAGCATGAT 68
QY 235 GGCAGCCTGGATATCTGGGCGCATTAAGAGAAAGCGGAAGCATGATCAGCATGTC 294
DB 69 GGAAGGCGCAAAATATGAGGCTATTAATCATCTGAAAGCATTAAGCATGATTAAGCATGTT 128
QY 295 CATAGTTTAAAGCAATATCTGATTCATTAAGTATGATCAAGTAACTTTTCTTTT 354
DB 129 GATTAACCTCAAACTTCAAGAGGTTACATACAGGATCAGCCGTAATTTTCTTA 188
QY 355 CAATCTGGGTAACTCTCACTTTTGAACAGATATGGGCACTAGCTGACATGAATAAT 414
DB 189 CAGTCAGGTCGCGGCGCGGCTTTTGAATATGGGCTTATCAGATCTGAACAG 248
QY 415 GATGGAAGATGATCAAGTGAAGTTTTCATAGCTATGAACTTTCAAACTGAAGCTA 474
DB 249 GATGGGAAGATGAGACCAAGATCTCTATAGCTATGAACTTTCATCAAGTTAAAGTTG 308
QY 475 CAAGATATCAGCTACC 491
DB 309 CAGGGCCCAACAGCTGCC 325

RESULT 4
US-09-216-003A-5
; Sequence 5, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-5

Query Match 27.6%; Score 137.8; DB 4; Length 531;
Best Local Similarity 64.7%; Pred. No. 1.1e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 175 GCGGTGATTAGCAAGTAAAGTAAACAGAACCATGGCTCACTTTCCAAACCTTTTGGT 234
DB 9 GCGTGAGAGCTGCAAGAAAGATCAGAGATCATGATGGCTCAGTTTCCACAGCATGAT 68
QY 235 GGCAGCCTGGATATCTGGGCGCATTAAGAGAAAGCGGAAGCATGATCAGCATGTC 294
DB 69 GGAAGGCGCAAAATATGAGGCTATTAATCATCTGAAAGCATTAAGCATGATTAAGCATGTT 128
QY 295 CATAGTTTAAAGCAATATCTGATTCATTAAGTATGATCAAGTAACTTTTCTTTT 354
DB 129 GATTAACCTCAAACTTCAAGAGGTTACATACAGGATCAGCCGTAATTTTCTTA 188
QY 355 CAATCTGGGTAACTCTCACTTTTGAACAGATATGGGCACTAGCTGACATGAATAAT 414
DB 189 CAGTCAGGTCGCGGCGCGGCTTTTGAATATGGGCTTATCAGATCTGAACAG 248
QY 415 GATGGAAGATGATCAAGTGAAGTTTTCATAGCTATGAACTTTCAAACTGAAGCTA 474
DB 249 GATGGGAAGATGAGACCAAGATCTCTATAGCTATGAACTTTCATCAAGTTAAAGTTG 308
QY 475 CAAGATATCAGCTACC 491
DB 309 CAGGGCCCAACAGCTGCC 325

RESULT 5
US-09-404-879A-72
; Sequence 72, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404.879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-72

Query Match 27.6%; Score 137.8; DB 4; Length 2017;
Best Local Similarity 64.7%; Pred. No. 2.3e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 175 GGGCTGATTAGCAAGTAAAGTAAACCAATGCTCAGTTTCCAAACCTTTTGGT 234
DB 9 GGCTGAGAGCTGCAAGAAAGAGTCAAGATCATGATGGCTCAGTTTCCCAAGCCATGAT 68
QY 235 GGCAGCTGATATCTGGGCCATTAAGTGAAGAAAGCAAGCATGATCAGAGTTT 294
DB 69 GGAGGGCCAAATATGTGGCTATTACATCTGAAGAACTACTAAGCATTAACAGTTT 128
QY 295 CATAGTTAAAGCCAAATATCTGGATTCTATTCTGGTATCAAGCTAAGAACTTTT 354
DB 129 GATTAACCTCAACCTTTCAGAGGTTACATTAAGGTGATCAAGCCGATCTTTTCCCTA 188
QY 355 CAATCTGGTTTACCTCAACCTGTTTTCACAGATATGGGCACTGATGATGAATAT 414
DB 189 CAGTAGGCTGCGCGGCCGCGTTTAACTGAAATATGGGCTTATCAGATCTGAACAAG 248
QY 415 GATGGAAGATGATCAAGTGAAGTTTCCATAGCTATGAATTAACACTGAAGCTA 474
DB 249 GATGGAAGATGATCAAGTGAAGTTTCTATAGCTATGAATTAACACTGAAGTTG 308
QY 475 CAAGGATATCAAGCTACC 491
DB 309 CAGGGCCCAACAGCTGCC 325

RESULT 6
US-09-338-933-72
; Sequence 72, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338.933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-72

Query Match 27.6%; Score 137.8; DB 4; Length 2017;

Best Local Similarity 64.7%; Pred. No. 2.3e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 175 GGGCTGATTAGCAAGTAAAGTAAACCAATGCTCAGTTTCCAAACCTTTTGGT 234
DB 9 GGCTGAGAGCTGCAAGAAAGAGTCAAGATCATGATGGCTCAGTTTCCCAAGCCATGAT 68
QY 235 GGCAGCTGATATCTGGGCCATTAAGTGAAGAAAGCAAGCATGATCAGAGTTT 294
DB 69 GGAGGGCCAAATATGTGGCTATTACATCTGAAGAACTACTAAGCATTAACAGTTT 128
QY 295 CATAGTTAAAGCCAAATATCTGGATTCTATTCTGGTATCAAGCTAAGAACTTTT 354
DB 129 GATTAACCTCAACCTTTCAGAGGTTACATTAAGGTGATCAAGCCGATCTTTTCCCTA 188
QY 355 CAATCTGGTTTACCTCAACCTGTTTTCACAGATATGGGCACTGATGATGAATAT 414
DB 189 CAGTAGGCTGCGCGGCCGCGTTTAACTGAAATATGGGCTTATCAGATCTGAACAAG 248
QY 415 GATGGAAGATGATCAAGTGAAGTTTCCATAGCTATGAATTAACACTGAAGCTA 474
DB 249 GATGGAAGATGATCAAGTGAAGTTTCTATAGCTATGAATTAACACTGAAGTTG 308
QY 475 CAAGGATATCAAGCTACC 491
DB 309 CAGGGCCCAACAGCTGCC 325

RESULT 7
US-09-215-681-72
; Sequence 72, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF OVARIAN CANCER
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215.681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-72

Query Match 27.6%; Score 137.8; DB 4; Length 2017;
Best Local Similarity 64.7%; Pred. No. 2.3e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 175 GGGCTGATTAGCAAGTAAAGTAAACCAATGCTCAGTTTCCAAACCTTTTGGT 234
DB 9 GGCTGAGAGCTGCAAGAAAGAGTCAAGATCATGATGGCTCAGTTTCCCAAGCCATGAT 68
QY 235 GGCAGCTGATATCTGGGCCATTAAGTGAAGAAAGCAAGCATGATCAGAGTTT 294
DB 69 GGAGGGCCAAATATGTGGCTATTACATCTGAAGAACTACTAAGCATTAACAGTTT 128
QY 295 CATAGTTAAAGCCAAATATCTGGATTCTATTCTGGTATCAAGCTAAGAACTTTT 354
DB 129 GATTAACCTCAACCTTTCAGAGGTTACATTAAGGTGATCAAGCCGATCTTTTCCCTA 188
QY 355 CAATCTGGTTTACCTCAACCTGTTTTCACAGATATGGGCACTGATGATGAATAT 414
DB 189 CAGTAGGCTGCGCGGCCGCGTTTAACTGAAATATGGGCTTATCAGATCTGAACAAG 248
QY 415 GATGGAAGATGATCAAGTGAAGTTTCCATAGCTATGAATTAACACTGAAGCTA 474
DB 249 GATGGAAGATGATCAAGTGAAGTTTCTATAGCTATGAATTAACACTGAAGTTG 308

QY 475 CAAGATATCAGCTACC 491
Db 309 CAGGCCAACAGCTGCC 325

RESULT 8

US-09-216-003A-72
; Sequence 72, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216.003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-72

Query Match 27.6%; Score 137.8; DB 4; Length 2017;
Best Local Similarity 64.7%; Pred. No. 2.3e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 175 GGGCTGATTAGCAGTAATAAGTAAACAGAACATGGCTCAGTTTCCAAACCTTTGGT 234
Db 9 GGCTGAGAGCTGCAAGAAAGATCAGAGATCATATGAGTCAAGTTTCCACAGCATGAAT 68
QY 235 GGAGGCTGATATCTGGGCACTACTAGAGAAAGCCAGAGATATCAGCAGTTT 234
Db 69 GGAGGCTGATATCTGGGCACTACTAGAGAAAGCCAGAGATATCAGCAGTTT 128
QY 295 CATAGTTTAAAGCAATATCTGATTCATTTACTGATGATCAAGCTGAAACTTTTTT 354
Db 129 GATTAACCTCAACCTTCAGAGAGTTTACATTAACAGGTATCAACCCGTTCTTTTCC 188
QY 355 CATCTGGGTTACTTCACTGTTTAAAGCAGATATGGGCACTAGCTGACATGAATAT 414
Db 189 CAGTCAAGGCTCTCGCCCGCTTTTAAAGTATGGGCTTATCAGATCTGAACAAG 248
QY 415 GATGAAGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 474
Db 249 GATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 308
QY 475 CAAGATATCAGCTACC 491
Db 309 CAGGCCAACAGCTGCC 325

RESULT 9

US-09-404-879A-60
; Sequence 60, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404.879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-404-879A-60

Query Match 26.7%; Score 133.6; DB 4; Length 480;
Best Local Similarity 66.9%; Pred. No. 2.2e-32;
Matches 190; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 208 ATGGCTCAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGCTATCATCTGAG 267
Db 14 ATGGCTCAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGCTATCATCTGAA 73
QY 268 GAAAGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
Db 74 GAACGATCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133
QY 328 GGTATCAAGCTTGAACCTTTTTCATCTGGGTTACTCAACCTGTTTAAACAG 387
Db 134 GGTATCAAGCTTGAACCTTTTTCATCTGGGTTACTCAACCTGTTTAAACAG 193
QY 388 ATATGGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
Db 194 ATATGGGCTTATCAATCTGAAACAGATGATGATGATGATGATGATGATGATGAT 253
QY 448 GCTATGAAACTTATCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 491
Db 254 GCTATGAAACTTATCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 297

RESULT 10

US-09-338-933-60
; Sequence 60, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338.933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-60

Query Match 26.7%; Score 133.6; DB 4; Length 480;
Best Local Similarity 66.9%; Pred. No. 2.2e-32;
Matches 190; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 208 ATGGCTCAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGCTATCATCTGAG 267
Db 14 ATGGCTCAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGCTATCATCTGAA 73
QY 268 GAAAGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
Db 74 GAACGATCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133
QY 328 GGTATCAAGCTTGAACCTTTTTCATCTGGGTTACTCAACCTGTTTAAACAG 387
Db 134 GGTATCAAGCTTGAACCTTTTTCATCTGGGTTACTCAACCTGTTTAAACAG 193
QY 388 ATATGGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
Db 194 ATATGGGCTTATCAATCTGAAACAGATGATGATGATGATGATGATGATGATGAT 253
QY 448 GCTATGAAACTTATCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 491
Db 254 GCTATGAAACTTATCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 297

RESULT 11

QY 345 CTTTTTTTCAATCTGGGTTTACCTCAACCTGTTTTCAGACAGATATGGGCACTGCTGA 404
DB 1780 GGAGATGGTGGCTCCCAAGTTTCCCAACAGAGTCTGGGCAAGATCTGGAAGTACCGCA 1839
QY 405 CATGAATTAATGATGGAAGATGATGATCAAGTGGAGTTTTCATGCTTAAGAACTTATCAA 464
DB 1840 CATTGACAGAGATGGCATGTTGATGACGAGAGTTTGGCTTCCCTGCCAACACCTTATCAA 1839
QY 465 ACTGAAGCTCAAGATATGCTACCGCTGCACT 500
DB 1900 AGTCAAGCTAGAGGGCATGAGCTCCCAAGTACT 1935

RESULT 14

US-09-312-762A-2
Sequence 2, Application US/09312762A
Patent No. 6552177
GENERAL INFORMATION:
APPLICANT: MIA HOROWITZ ET AL.
TITLE OF INVENTION: EH DOMAIN CONTAINING GENES AND PROTEINS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
STREET: 2001 Jefferson Davis Highway, Suite 207
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
COMPUTER: Twinhead* Slimote-890TX
OPERATING SYSTEM: MS DOS version 6.2,
OPERATING SYSTEM: Windows version 3.11
SOFTWARE: Word for Windows version 2.0 converted to
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312,762A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/026,898
FILING DATE: 20 FEB 1998
ATTORNEY/AGENT INFORMATION:
NAME: Friedman, Mark M.
REGISTRATION NUMBER: 33,883
REFERENCE/DOCKET NUMBER: 916/10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 972-3-5625553
TELEFAX: 972-3-5625554
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3348
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-312-762A-2

Query Match 13.8%; Score 69.2; DB 4; Length 3348;
Best Local Similarity 53.8%; Pred. No. 1.4e-11;
Matches 143; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 235 GGCAGCCTGATATCTGGGCACTAATCTGTAGAGAAAGCGAAGCATGATGACAGTTTC 294
DB 1379 GGCATTGATGATGTTAGTGGTATGTTGGCAAGACAGCCCACTATATGATGATCTTC 1438
QY 295 CATAGTTTAAAGCAATATCTGATTCATTACTGTGATCAAGCTTGAACCTTTTTCCTT 354
DB 1439 TACACACTGTCTCTGTCACAGGCAAGATCAAGGTGCTAATGCCAAGAGAGATGCTG 1498
QY 355 CAATCTGGGTTACCTCAACCTTTTTCAGACAGATATGCGGCACTAGCTACATGAATAT 414

DB 1499 AAGTCCAGAGTCCCAACACAGTCTGGGGAAGATCTGGAAGTTGCGAGATGACAG 1558
QY 415 GATGGAAGATGATCAAGTGGAGTTTTCATGATGAACTTATGAACTTATCAAGCTTA 474
DB 1559 GATGGCTGTGATGATGACGAGAGTTTCCCTGCCAACACCTTATCAAGTGAAGCTA 1618
QY 475 CAAGATATGACGATCCTGCACT 500
DB 1619 GAGGCCACGAGCTGCCCTGACT 1644

RESULT 15

US-09-312-762A-3
Sequence 3, Application US/09312762A
Patent No. 6552177
GENERAL INFORMATION:
APPLICANT: MIA HOROWITZ ET AL.
TITLE OF INVENTION: EH DOMAIN CONTAINING GENES AND PROTEINS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
STREET: 2001 Jefferson Davis Highway, Suite 207
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
COMPUTER: Twinhead* Slimote-890TX
OPERATING SYSTEM: MS DOS version 6.2,
OPERATING SYSTEM: Windows version 3.11
SOFTWARE: Word for Windows version 2.0 converted to
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312,762A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/026,898
FILING DATE: 20 FEB 1998
ATTORNEY/AGENT INFORMATION:
NAME: Friedman, Mark M.
REGISTRATION NUMBER: 33,883
REFERENCE/DOCKET NUMBER: 916/10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 972-3-5625553
TELEFAX: 972-3-5625554
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 14707
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-312-762A-3

Query Match 13.8%; Score 69.2; DB 4; Length 14707;
Best Local Similarity 53.8%; Pred. No. 3.3e-11;
Matches 143; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 235 GGCAGCCTGATATCTGGGCACTAATCTGTAGAGAAAGCGAAGCATGATGACAGTTTC 294
DB 9100 GGCATTGATGATGTTAGTGGTATGTTGGCAAGACAGCCCACTATATGATGATCTTC 9159
QY 295 CATAGTTTAAAGCAATATCTGATTCATTACTGTGATCAAGCTTGAACCTTTTTCCTT 354
DB 9160 TACACACTGTCTCTGTCACAGGCAAGATCAAGGTGCTAATGCCAAGAGAGATGCTG 9219
QY 355 CAATCTGGGTTACCTCAACCTTTTTCAGACAGATATGCGGCACTAGCTACATGAATAT 414
DB 9220 AAGTCCAGAGTCCCAACACAGTCTGGGGAAGATCTGGAAGTTGCAAGATGCGCAAG 9279
QY 415 GATGGAAGATGATCAAGTGGAGTTTTCATAGCTATGAACTTATCAAGTGAAGCTTA 474

Db	9280	GATGCGCTGCTGAGATGACGAGGAGTTGCCCTGGCCACACACCTTATCAGGTGAGCTA	9339
Qy	475	CAAGGATATCAGCTACCCCTCTGCACT	500
Db	9340	GAGGCCACGAGCTGCCGCTGACT	9365

Search completed: July 1, 2004, 13:02:56
 Job time : 93 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2004, 11:34:21 ; Search time 363 Seconds
(without alignments)
6644.914 Million cell updates/sec

Title: US-09-720-934-1_COPY_1_500
Perfect score: 500
Sequence: 1 caaagaatccgggtacgg.....taccgtaccctctcact 500

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3163042 seqs, 2412103800 residues
Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PTCTIS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	486	97.2	3466	16	US-10-158-057-33 Sequence 33, Appl
2	444.4	88.9	3319	11	US-09-764-875-88 Sequence 89, Appl
3	345	69.0	568	10	US-09-764-881-55 Sequence 55, Appl
4	345	69.0	568	11	US-09-764-875-404 Sequence 404, App
5	345	69.0	568	13	US-09-764-881-55 Sequence 55, Appl
6	345	69.0	568	16	US-10-242-747-55 Sequence 55, Appl
7	345	69.0	568	16	US-10-158-057-127 Sequence 127, App
8	152.8	30.6	967	10	US-09-764-881-50 Sequence 50, Appl
9	152.8	30.6	967	11	US-09-764-875-411 Sequence 411, App
10	152.8	30.6	967	13	US-09-764-881-50 Sequence 50, Appl
11	152.8	30.6	967	16	US-10-242-747-50 Sequence 40, Appl
12	152.8	30.6	967	16	US-10-158-057-40 Sequence 15, Appl
13	152.8	30.6	5828	13	US-10-398-895A-15 Sequence 5, Appl
14	137.8	27.6	531	9	US-09-884-441-5

15	137.8	27.6	531	10	US-09-907-969-5	Sequence 5, Appl
16	137.8	27.6	531	10	US-09-827-271-5	Sequence 5, Appl
17	137.8	27.6	531	15	US-10-198-053-5	Sequence 5, Appl
18	137.8	27.6	2017	9	US-09-884-441-72	Sequence 72, Appl
19	137.8	27.6	2017	10	US-09-907-969-72	Sequence 72, Appl
20	137.8	27.6	2017	9	US-09-827-271-72	Sequence 72, Appl
21	137.8	27.6	2017	15	US-10-198-053-72	Sequence 72, Appl
22	133.6	26.7	480	9	US-09-884-441-60	Sequence 60, Appl
23	133.6	26.7	480	10	US-09-907-969-60	Sequence 60, Appl
24	133.6	26.7	480	10	US-09-827-271-60	Sequence 60, Appl
25	133.6	26.7	480	15	US-10-198-053-60	Sequence 60, Appl
26	122.4	24.5	462	10	US-09-918-995-21728	Sequence 21728, A
27	109	21.8	292	9	US-09-864-761-20261	Sequence 20261, A
28	109	21.8	310	9	US-09-864-761-21373	Sequence 21373, A
29	109	21.8	310	9	US-09-864-761-19751	Sequence 19751, A
30	109	21.8	310	9	US-09-864-761-19759	Sequence 19759, A
31	109	21.8	463	9	US-09-864-761-2978	Sequence 2978, Ap
32	109	21.8	465	9	US-09-864-761-2970	Sequence 2970, Ap
33	109	21.8	466	9	US-09-864-761-3491	Sequence 3491, Ap
34	109	21.8	466	9	US-09-864-761-4631	Sequence 4631, Ap
35	107	21.4	307	9	US-09-864-761-30533	Sequence 30533, A
36	79.2	15.8	3635	9	US-09-312-762A-6	Sequence 6, Appl
37	71.4	14.3	605	12	US-10-152-319A-1126	Sequence 1126, Ap
38	69.2	13.8	3348	9	US-09-312-762A-2	Sequence 2, Appl
39	69.2	13.8	14707	9	US-09-312-762A-3	Sequence 3, Appl
40	68.4	13.7	1021	11	US-09-764-875-96	Sequence 96, Appl
41	61.8	12.4	2259	16	US-10-369-493-26020	Sequence 26020, A
42	59.4	11.9	2442	9	US-09-964-899-24	Sequence 24, Appl
43	58	11.6	551	15	US-10-029-386-25205	Sequence 25205, A
44	55.8	11.2	590	9	US-09-879-536-211	Sequence 211, App
45	55.8	11.2	2520	13	US-10-302-172-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-10-158-057-33
Sequence 33, Application US/10158057
Publication No. US20040014039A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P1205051
CURRENT APPLICATION NUMBER: US/10/158,057
CURRENT FILING DATE: 2002-06-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 384
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 33
LENGTH: 3466
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (3194)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (3465)
OTHER INFORMATION: n equals a,t,g, or c
US-10-158-057-33

Query Match 97.2%; Score 486; DB 16; Length 3466;
Best Local Similarity 100.0%; Pred. No. 1.7e-147;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GGAAGGCGGCTGCGAGGAATCCGAGCGGCTCCGAGCGAGAGGAGCGGCG 74
DB 84 GAGCGGCGGCTGCGAGGAATCCGAGCGGCTCCGAGCGAGAGGAGCGGCG 143
QY 75 GGAAGGCTGCGCGGCGGCTGCGGCTCCGAGCGGCTGAGCGGCGACTGA 134

Db	144	GGAGATGATGCGGGGCTCGCTCTGGCTCCTCCACGGGCGGTGAGGGCACTGA	203
QY	135	TTTGTTCCTCGGGGCGGCAAGCGGACCCGCGAGATGAGGGGTGCATTAGCAAGTTAA	194
Db	204	TTTGTCCCTGGGGCGGCAAGCGGAGCCCGCCGGAGATGAGGGGTGCATTAGCAAGTTAA	263
QY	195	AAGTAAACGAAACCATGGCTCAGTTTCCAAACCTTTTGGGGAGGCTGGATATCTGGGC	254
Db	264	AAGTAAACGAAACCATGGCTCAGTTTCCAAACCTTTTGGGGAGGCTGGATATCTGGGC	323
QY	255	CATTAACGTGAGGAAAGACGAAGCATGATCAGCAGTTCCATAGTTTAAAGCCAATATC	314
Db	324	CATACTGTAGAGGAAAGACGAAGCATGATCAGCAGTTCCATAGTTTAAAGCCAATATC	383
QY	315	TGGATTCATTTACTGGTGCATCAGCTGAAAACTTTTTTTCAATCTGGGTTACCTCAACC	374
Db	384	TGGATTCATTTACTGGTGCATCAGCTGAAAACTTTTTTTCAATCTGGGTTACCTCAACC	443
QY	375	TGTTTACACACGATATGGGGCACTAGCTGCATCAGTAAGTAATATGGAAGAAATGGATCAAGT	434
Db	444	TGTTTACACACGATATGGGGCACTAGCTGCATCAGTAAGTAATATGGAAGAAATGGATCAAGT	503
QY	435	GGAGTTTCCATAGCTATGAACCTTATCAAACTGAAGCTACAAGGATATACAGTACCCCTC	494
Db	504	GGAGTTTCCATAGCTATGAACCTTATCAAACTGAAGCTACAAGGATATACAGTACCCCTC	563
QY	495	TGCACCT 500	
Db	564	TGCACCT 569	

```

RESULT 2
US-09-764-875-88
: Sequence 88, Application US/09764475
: Publication No. US20040018969A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P1202
: CURRENT APPLICATION NUMBER: US/09/764,875
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed -
: NUMBER OF SEQ ID NOS: 1249
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 88
: LENGTH: 3319
: TYPE: DNA
: ORGANISM: Homo sapiens
: IS-09-764-875-88

```

Query Match	88.9%	Score 444.4	DB 11	Length 3319
Best Local Similarity	99.8%	Pseq No. 6.6e-134		
Matches 445	Conservative 0	Mismatches 1	Indels 0	Gaps 0

[illegible]

QY 355 CAATCTGGGTTTACCTCAACTGTTTATTAGACAGATATGGGCACTAGCTACATGAATTAAT 414

Db 306 CAATTGGGTTTACCTCAACTGTTTATTAGACAGATATGGGCACTAGCTACATGAATTAAT 365

QY 415 GATGAGAGATGATCAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTGAAGCTA 474

Db 366 GATGAGAGATGATCAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTGAAGCTA 425

QY 475 CAAGGATATCAGCTACCTCTGCTCACT 500

Db 426 CAAGGATATCAGCTACCTCTGCTCACT 451

```

RESULT 3
US-09-764-881-55
; Sequence 55, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55

```

```

? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (461)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (536)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (556)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (562)
? OTHER INFORMATION: n equals a,t,g, or c
? OS-09-764-881-55

```

Query Match	69.0%	Score 345;	DB 10;	Length 568;
Best Local Similarity	99.2%	Pred. NO. 7.2e-102;		
Matches 368;	Conservative 0;	Mismatches 0;	Indels 3;	Gaps 2;

QY	13	ACTGATTTGTCCTCGTGGGCGGAGCGGCGAACCCGCGCGGATCATAGGCTGTGATTTACAA	189
Db	1	ACTGATTTTTCCTCGTGGGCGGCGA-CGGGACCCCGCGGATCATAGGCTGTGATTTACAA	59
QY	190	GGTAAAGTAAGAAACCATGGCTCAGTTTCCAACTCTTTTGGTGGGACGCTGGATATC	249
Db	60	GGTAAAGTAAGAAACCATGGCTCAGTTTCCAACTCTTTTGGTGGGACGCTGGATATC	119
QY	250	TGGGCGCATTAAGTGTAGAGGAAAGAGCGGAGCATGATCAGCAGTTTCCATAGTTTAAAGCCA	309
Db	120	TGGGCGCATTAAGTGTAGAGGAAAGAGCGGAGCATGATCAGCAGTTTCCATAGTTTAAAGCCA	177
QY	310	AATATCTGGAATTCATTAACGTGGATCAAGCTTGAAGAACTTTTTTTTCAATCTGGGTTAAGT	369
Db	178	AATATCTGGAATTCATTAACGTGGATCAAGCTTGAAGAACTTTTTTTTCAATCTGGGTTAAGT	237
QY	370	CAACTGTGTTTTAGCACAGATATGGGCACTAGCTGACATGATATATGATGGAAAGATGAT	429
Db	238	CAACTGTGTTTTAGCACAGATATGGGCACTAGCTGACATGATATATGATGGAAAGATGAT	297
QY	430	CAAGTGGAGTTTCCATAGCTATGATTAACCTTATCAAACCTGAAGCTACAGGATATCAGCTA	489
Db	298	CAAGTGGAGTTTCCATAGCTATGATTAACCTTATCAAACCTGAAGCTACAGGATATCAGCTA	357

QY 490 CCCTCTGCACT 500
Db 358 CCCTCTGCACT 368

RESULT 4

US-09-764-875-404
; Sequence 404, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 404
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-404

Query Match 69.0%; Score 345; DB 11; Length 568;

Best Local Similarity 99.2%; Pred. No. 7.2e-102;
Matches 368; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 130 ACTGATTTGTCCTGGGGGCGGAGCGCGACCCCGCGAGATGAGGGGTGATTAAGCA 189
Db 1 ACTGATTTGTCCTGGGGGCGGAGCGCGACCCCGCGAGATGAGGGGTGATTAAGCA 59
QY 190 GGTAAAGTAACAGAACCATGCTGAGTTTCCACACCTTTGGTGGAGCCTGGATATC 249
Db 60 GGTAAAGTAACAGAACCATGCTGAGTTTCCACACCTTTGGTGGAGCCTGGATATC 119
QY 250 TGGGCCATPACTGTAGAGGAAAGAGGAGCATGATCAGACGTTCCATGTTAAAGCA 309
Db 120 TGGGCCATPACTGTAGAGGAAAGAGGAGCATGATCAGACGTTCCATGTTAAAGCA 177
QY 310 ATATCTGATTCATTACTGCTGATCAAGTAACTTTTTCATCTGGTTACT 369
Db 178 ATATCTGATTCATTACTGCTGATCAAGTAACTTTTTCATCTGGTTACT 237
QY 370 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 429
Db 238 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 297
QY 430 CAAGTGAAGTTTCCATAGCTATGAACCTTATCAACTGAAGCTACAGATATCAGCTA 489
Db 298 CAAGTGAAGTTTCCATAGCTATGAACCTTATCAACTGAAGCTACAGATATCAGCTA 357
QY 490 CCCTCTGCACT 500
Db 358 CCCTCTGCACT 368

RESULT 5
US-09-764-881-55

; Sequence 55, Application US/09764881
; Publication No. US20020086821A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-55

Query Match 69.0%; Score 345; DB 13; Length 568;

Best Local Similarity 99.2%; Pred. No. 7.2e-102;
Matches 368; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 130 ACTGATTTGTCCTGGGGGCGGAGCGCGACCCCGCGAGATGAGGGGTGATTAAGCA 189
Db 1 ACTGATTTGTCCTGGGGGCGGAGCGCGACCCCGCGAGATGAGGGGTGATTAAGCA 59
QY 190 GGTAAAGTAACAGAACCATGCTGAGTTTCCACACCTTTGGTGGAGCCTGGATATC 249
Db 60 GGTAAAGTAACAGAACCATGCTGAGTTTCCACACCTTTGGTGGAGCCTGGATATC 119
QY 250 TGGGCCATPACTGTAGAGGAAAGAGGAGCATGATCAGACGTTCCATGTTAAAGCA 309
Db 120 TGGGCCATPACTGTAGAGGAAAGAGGAGCATGATCAGACGTTCCATGTTAAAGCA 177
QY 310 ATATCTGATTCATTACTGCTGATCAAGTAACTTTTTCATCTGGTTACT 369
Db 178 ATATCTGATTCATTACTGCTGATCAAGTAACTTTTTCATCTGGTTACT 237
QY 370 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 429
Db 238 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 297
QY 430 CAAGTGAAGTTTCCATAGCTATGAACCTTATCAACTGAAGCTACAGATATCAGCTA 489
Db 298 CAAGTGAAGTTTCCATAGCTATGAACCTTATCAACTGAAGCTACAGATATCAGCTA 357
QY 490 CCCTCTGCACT 500
Db 358 CCCTCTGCACT 368

RESULT 6

US-10-242-747-55
; Sequence 55, Application US/10242747
; Publication No. US20040005577A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207C1
; CURRENT APPLICATION NUMBER: US/10/242,747
; CURRENT FILING DATE: 2002-09-13

PRIOR APPLICATION NUMBER: 09/764,881
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data
NUMBER OF SEQ ID NOS: 192
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 55
LENGTH: 568
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (481)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (536)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (556)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (562)
OTHER INFORMATION: n equals a,t,g, or c
US-10-242-747-55

Query Match 69.0%; Score 345; DB 16; Length 568;
Best Local Similarity 99.2%; Pred. No. 7, 2e-102;
Matches 368; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 130 ACTGATTTCCTCGGGGCGGACCGCGGACCGCGGAGAGTGAAGCGCTGATTAGCAA 189
DB 1 ACTGATTTCCTCGGGGCGGACCGCGGACCGCGGAGAGTGAAGCGCTGATTAGCAA 59
QY 190 GGTAAAGTAACGAACCATGGCTCAGTTTCCAAACCTTTGGTGGCAGCCTGGATATC 249
DB 60 GGTAAAGTAACGAACCATGGCTCAGTTTCCAAACCTTTGGTGGCAGCCTGGATATC 119
QY 250 TGGGCCATTAAGTGAAGGAAAGAGGAGCATGATCAGAGTTCATGTTTAAAGCCA 309
DB 120 TGGGCCATTAAGTGAAGGAAAGAGGAGCATGATCAGAGTTCATGTTTAAAGCCA 177
QY 310 ATATCTGATTAATCTGATGATCAAGCTAGAACTTTTTTTCATCTGGGTTACCT 369
DB 178 ATATCTGATTAATCTGATGATCAAGCTAGAACTTTTTTTCATCTGGGTTACCT 237
QY 370 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGAATATGATGAAGATGAT 429
DB 238 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGAATATGATGAAGATGAT 297
QY 430 CAAGTGAAGTTTCCATGATGATGAAGCTTATCAACTGATCAAGCTATCAAGATAT 489
DB 298 CAAGTGAAGTTTCCATGATGATGAAGCTTATCAACTGATCAAGCTATCAAGATAT 357

QY 490 CCCTCTGACT 500
DB 358 CCCTCTGACT 368

RESULT 7
US-10-158-057-127
Sequence 127, Application US/10158057
Publication No. US20040014039A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PZ05051
CURRENT APPLICATION NUMBER: US/10/158,057
Prior Filing Date: 2002-06-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 384
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 127
LENGTH: 568
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (481)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (536)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (556)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (562)
OTHER INFORMATION: n equals a,t,g, or c
US-10-158-057-127

Query Match 69.0%; Score 345; DB 16; Length 568;
Best Local Similarity 99.2%; Pred. No. 7, 2e-102;
Matches 368; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 130 ACTGATTTCCTCGGGGCGGACCGCGGACCGCGGAGAGTGAAGCGCTGATTAGCAA 189
DB 1 ACTGATTTCCTCGGGGCGGACCGCGGACCGCGGAGAGTGAAGCGCTGATTAGCAA 59
QY 190 GGTAAAGTAACGAACCATGGCTCAGTTTCCAAACCTTTGGTGGCAGCCTGGATATC 249
DB 60 GGTAAAGTAACGAACCATGGCTCAGTTTCCAAACCTTTGGTGGCAGCCTGGATATC 119
QY 250 TGGGCCATTAAGTGAAGGAAAGAGGAGCATGATCAGAGTTCATGTTTAAAGCCA 309
DB 120 TGGGCCATTAAGTGAAGGAAAGAGGAGCATGATCAGAGTTCATGTTTAAAGCCA 177
QY 310 ATATCTGATTAATCTGATGATCAAGCTAGAACTTTTTTTCATCTGGGTTACCT 369
DB 178 ATATCTGATTAATCTGATGATCAAGCTAGAACTTTTTTTCATCTGGGTTACCT 237
QY 370 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGAATATGATGAAGATGAT 429
DB 238 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGAATATGATGAAGATGAT 297
QY 430 CAAGTGAAGTTTCCATGATGATGAAGCTTATCAACTGATCAAGCTATCAAGATAT 489
DB 298 CAAGTGAAGTTTCCATGATGATGAAGCTTATCAACTGATCAAGCTATCAAGATAT 357

```
RESULT 8
US-09-764-881-50
; Sequence 50, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 967
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (659)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (697)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (715)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (869)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (879)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-50

Query Match      30.6%; Score 152.8; DB 10; Length 967;
Best Local Similarity 67.5%; Pred. No. 7.8e-39;
Matches 208; Conservative 4; Mismatches 96; Indels 0; Gaps 0;

QY 184 TAGCAAGGTAAAGTAACAGAACCATGCTCAGTTCCAGACCTTTGGTGGCAGCCTG 243
DB 218 TAGCAAGGAAAACTCAGAACCATGATGCTCAGTTCCAGACGTATGAATGGAGGCCA 277
QY 244 GATATCTGGCCATTAATCTGTAAGAGAAAGGAGCATGATCAGCACTTCATGTTTA 303
DB 278 AACATGKGTATTAATCTCTGAGAACGTACTAAGCATGACAGCAGTTGATTAACCTC 337
QY 304 AAGCAATATCTGATTCATTAATCTGATCAAGCTAAGAACTTTTTCATCAATCTGGG 363
DB 338 AAACCTTGAAGAGTTTACATTAACAGGTGATCAAGCAGTAATTTTCTTACAAATCAGGT 397
QY 364 TTACCTCAACCTGTTTTAGCAGATATGCGCAGTACAGTGAATGATGATGAGAA 423
DB 398 CTGCCGGCCCTGTTTAGCTGGAATATGGGCTTATCAGACCTAAACAMGKATGGAG 457
QY 424 ATGATCAAGTGAAGTTTCCATAGCTATGAACCTTATCAAACTTAAGCTACAGATAT 483
DB 458 ATGATTCAGCAAGAGTTCTCATATAGTAAGAACTCATCAAACTGAAGCTTCAAGCCAA 517
QY 484 CAGCTACC 491
DB 518 CAGTTGCC 525
```

```
RESULT 9
US-09-764-875-411
; Sequence 411, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT202
; CURRENT APPLICATION NUMBER: US/09/764,875

RESULT 10
US-09-764-881-50
; Sequence 50, Application US/09764881
; Publication No. US20020086821A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 967
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 411
; LENGTH: 967
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (659)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (697)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (715)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (869)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (879)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-411

Query Match      30.6%; Score 152.8; DB 11; Length 967;
Best Local Similarity 67.5%; Pred. No. 7.8e-39;
Matches 208; Conservative 4; Mismatches 96; Indels 0; Gaps 0;

QY 184 TAGCAAGGTAAAGTAACAGAACCATGCTCAGTTCCAGACCTTTGGTGGCAGCCTG 243
DB 218 TAGCAAGGAAAACTCAGAACCATGATGCTCAGTTCCAGACGTATGAATGGAGGCCA 277
QY 244 GATATCTGGCCATTAATCTGTAAGAGAAAGGAGCATGATCAGCACTTCATGTTTA 303
DB 278 AACATGKGTATTAATCTCTGAGAACGTACTAAGCATGACAGCAGTTGATTAACCTC 337
QY 304 AAGCAATATCTGATTCATTAATCTGATCAAGCTAAGAACTTTTTCATCAATCTGGG 363
DB 338 AAACCTTGAAGAGTTTACATTAACAGGTGATCAAGCAGTAATTTTCTTACAAATCAGGT 397
QY 364 TTACCTCAACCTGTTTTAGCAGATATGCGCAGTACAGTGAATGATGATGAGAA 423
DB 398 CTGCCGGCCCTGTTTAGCTGGAATATGGGCTTATCAGACCTAAACAMGKATGGAG 457
QY 424 ATGATCAAGTGAAGTTTCCATAGCTATGAACCTTATCAAACTTAAGCTACAGATAT 483
DB 458 ATGATTCAGCAAGAGTTCTCATATAGTAAGAACTCATCAAACTGAAGCTTCAAGCCAA 517
QY 484 CAGCTACC 491
DB 518 CAGTTGCC 525
```

```

: NAME/KEY: SITE
: LOCATION: (659)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (697)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (715)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (869)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (879)
: OTHER INFORMATION: n equals a,t,g, or c
: OS-09-764-881-50

```

Query Match	30.6%;	Score 152.8;	DB 13;	Length 967;
Best Local Similarity	67.5%;	Pred. No. 7.8e-39;		
Matches 208; Conservative	4;	Mismatches 96;	Indels 0;	Gaps 0;

Qy	18	TACCAAGTAAAGTAACAGAACCTAGGTGAGTTTCCAAACCTTTTGGGGGAGCTG	2438
Db	218	TACCAAGAAAACCTCAGAACCATGATGGCTCAGTTTCCCAAGCTAATGATGAGGGCCA	2777
Qy	244	GATATCTGGCCATPACTGTABAGGAAAGCGAAGCATGATCAGAGTTCCATATGTTTA	3030
Db	278	AACATGTGGCTATTATACCTCTAAGAAAGCTAATGACATGCAAGGCTTTGATTAACCTC	3337
Qy	304	AAGCCATATATCTGATTGATTACTGGTGTATCAGCTAATAAATTCTTTTTCATCTCGG	3653
Db	338	AAACCTTCAGAGGTTACATTAACAGTATCAGACGATAATTTTTCCTACATTCAGGT	3937
Qy	364	TTAAGCTAAACCTGTTTACGATAGATATGGGCACTAGTCATAGTAATATGATGAGAA	4233
Db	398	CTGGCGGCCCCGTTTATAGCTAATAATATGGGCTTTATAGACCTTAACAGKATGGGAG	4577
Qy	424	ATGATCAAGTGGAGTTTTCATAGCTATGAAACTTATCAACAGTAAGCTAACAAGATAT	483
Db	458	ATGATCAGCAAGAGTTTCCTCATAGCTATGAAACTCATCAACTGAAGCTTCAAGGCCAA	517
Qy	484	CAGCTAAC	491
Db	518	CAGTTGCC	525

RESULT 11
US-10-242-747-50
Sequence 50, Application US/10242747
Publication No. US20040005577A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PRZ07C1
CURRENT APPLICATION NUMBER: US/10/242,747
CURRENT FILING DATE: 2002-09-13
PRIOR APPLICATION NUMBER: 09/784,881
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/119,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447

```

; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50

```

```

1 ORGANISM: Homo sapiens
2 FEATURE:
3 NAME/KEY: misc_feature
4 LOCATION: (659)
5 OTHER_INFORMATION: n equals a,t,g, or c
6 FEATURE:
7 NAME/KEY: misc_feature
8 LOCATION: (697)
9 OTHER_INFORMATION: n equals a,t,g, or c
10 FEATURE:
11 NAME/KEY: misc_feature
12 LOCATION: (715)
13 OTHER_INFORMATION: n equals a,t,g, or c
14 FEATURE:
15 NAME/KEY: misc_feature
16 LOCATION: (869)
17 OTHER_INFORMATION: n equals a,t,g, or c
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: (879)
21 OTHER_INFORMATION: n equals a,t,g, or c
22 IS-10-242..747-50

```

Query Match	30.6%;	Score 152.8;	DB 16;	Length 967;
Best Local Similarity	67.5%;	Pred. No. 7.8e-39;		
Matches 208; Conservative	4;	Mismatches 96;	Indels 0;	Gaps 0;

Qy	18	TAGCAGGTAAAGAATACGAAACCACTGGCTAGTTTCCAAACCTTTTGGGCGAGCTG	243
Db	218	TAGCAAGGAAAACTCGAGCAGCATGATGCTAGTTTCCACAGCTTAGATGAGGGCCA	277
Qy	244	GATATCTGGGCGATTACTGTAGAGGAAGAAGCGAAGCATGATCGAGTTCCATGATTTA	303
Db	278	AACATCTGGCTATTACCTCTGAGAGAGCGTACTAAGCATGACAGGCGATTTGATTAACCTC	337
Qy	304	AAGCCAAATCTGAGATTCATTACTGTGATCAAGCTAGGAACTTTTTTTTCATCTGGG	363
Db	338	AAACCTTCAGGAGGTTAATATACAGGTATACAGCATTAATTTTTCCTACATCTCAGGT	397
Qy	364	TTAAGCTCAACCTGTTTTAGCACAGATATGGGCACTAGCTGACATGAATTAATGATGAGAA	423
Db	398	CTGCGGGCCCCCTGTTTTAGCTGAATAATGGGCTTTATCAGCTAAACAMCKATGGGAG	457
Qy	424	ATGGAATCAAGTGAGCTTTCCATAGCTGTGAACCTTATCAACAGTGAAGCTCAAGATAT	483
Db	458	ATGGAATGAGCAAGAGTTCTCATAGCTATGAATCATCAAACTGAAGCTTCAGAGCCAA	517
Qy	484	CAGCTACC	491
Db	518	CAGTTGCC	525

```

RESULT 12
US-10-158-057-40
; Sequence 40, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P2505C1
; CURRENT APPLICATION NUMBER: US/10/158,057
; CURRENT FILING DATE: 2002-06-12
; Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 384

```

Query Match	30.6%	Score 152.8	DB 16	Length 967
Best Local Similarity	67.5%	Pred. NO. 7.8e-39		
Matches 208; Conservative	4;	Mismatches 96;	Indels 0;	Gaps 0

QY	184	TAGCAAGGTAAAGTAAAGTAAAGAACGAAACATGGCTCAGTTTCCAAACACCTTTGGTGGCAGCCTG	243
Db	218	TAGCAAGGAAAACATCAGGACCATGATGGCTCAGTTTCCACAGCTATGAATGAAGGACA	277
QY	244	GATATCTGGGCGCATPACTCTGTAGAGAAAGAGGAAGCATGATACGACATCTCATGTGTTA	303
Db	278	AAACATGTGKGTATTAATCTCTGAAAGAAAGTAAAGACATGACAGGACGATTTGATTAACCTC	337
QY	304	AAGCCAAATATCTGATTCATTACTGGATCATCAAGCTAGAAAATTTTTTTTCAATCTGGG	363
Db	338	AAACCTTCAGGAGGTTATACATAACAGGTGATCAAGCAGTAATTTTTTCTCAATCAGGT	397
QY	364	TTACTCTCAACTGTTTGTAGCAGACATATGGGACATAGCTGACATGAATATATGATGGAA	423
Db	398	CTGCCGGCCCTGTTTGTAGCTGAATATATGGGCTTTATACAGCTTAAACCMGATGGGABG	457
QY	424	ATGATCTCAAGTGAGATTTTCCATACCTATGAACTTATCAAACTGAAGCTACMAAGATAT	483
Db	458	ATGATCTAGCAAGAAATTTCTCCATACCTATGAAACATCATCAAACTGAAGCTTCAAGGCCA	517
QY	484	CAGCTAAC	491
Db	518	CAGTTGCC	525

```

RESULT 13
US-10-398-885A-15
; Sequence 15, Application US/10398885A
; Publication No. US20040053282A1
; GENERAL INFORMATION:
; APPLICANT: Sugita, Yuji
; APPLICANT: Hashida, Ryotachi
; APPLICANT: Ogawa, Kaoru
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Obayashi, Masaya
; APPLICANT: Saito, Hirohisa
; APPLICANT: Takahashi, Eiki
; TITLE OF INVENTION: Method of Testing For Allergic Diseases
; FILE REFERENCE: SHIMURU-07907
; CURRENT APPLICATION NUMBER: US/10/398,885A
; CURRENT FILING DATE: 2003-08-11
; PRIORITY APPLICATION NUMBER: PCT/JP01/08937

```

Query Match	30.6%	Score 152.8	DB 13	Length 5828
Best Local Similarity	68.5%	Pred. No. 2,5e-38		
Matches 211	Conservative 0	Mismatches 97	Indels 0	Gaps 0

QY	184	TAGCAGAGTAAATACTACAGAACCATGGCTAGTTTCCAAACCTTTGGTGGCAGGCTG	243
Db	19	TATCAAGSAAAACTCAGGACCATGATGGCTCAGTTTCCACAGCTATGAATGAGGGCCA	78
QY	244	GATATCTGGGSCCTAACTGTAGAGSAAAAGACCAAGCATGATTCAGCACTTCCATAGTTTA	303
Db	79	AAACATGGGGCTATTACCTCTGAAGAAGAGTCTAAGCATGACAGGCACTTTGATACCTTC	138
QY	304	AAGCAATATCTGGATTCATCTACTGATGATCAAGCTAGAAACTTTTCTTTTCAATCTGGG	363
Db	139	AAACCTTCAGSAGGTTTCATTAACAGAGTGATCAGACAGCTAAATTTTCTCCATCAATCAGST	198
QY	364	TTACTTCAACTGCTTTTTAGCACACATATNGGSCACTAGCTGACATGAATAATAGATGAGA	423
Db	199	CTGGCGGCCCTGTTTATGCTGAAATATGGGCTTTATAGACCTAAACAGGATGGGAG	258
QY	424	ATGATCAACAGTGAAGTTTCCATAGCTATGAACTTATCAAACTGAACCTTCAAGATAT	483
Db	259	ATGATCAAGCAGAGGTTCTCCATGACATGAACTCAATCAAACTGAACCTTCAAGGCCAA	318
QY	484	CAGCTACC	491
Db	319	CAGTTGCC	326

```

RESULT 14
US-09-884-441-5
; Sequence 5, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darlick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.463CT
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-5

```

Query	175	GGCGTCGATTGCGAAGCTAAAGTACAGCAACGACGCTCGATTCCTCCACACCTTTGGT	234
Db	9	GGCTGAGCGCTCGAAGAAAGTCAAGCTACGGATTCATGATGGCTCATCTTCTCCACAGCATGAAT	68
Query Match	27.6%	Score 137.8	DB 9; Length 531;
Best Local Similarity	64.7%;	Pred. No. 4.3e-34;	
Matches	205; Conservative	0; Mismatches 112;	Indels 0; Gaps 0

QY	235	GGCAGCCCTGAGATCTCGGGCCCACTA	CTGTGAGGAGAAAGAGCCAGATATCAGACGTT	294
Db	69	GGAGGCCCCAAATATGTGGGCTTTCAT	CTGTGAAGACGACTTGAAGATGATTAACGTTT	128
QY	295	CATAGTTTAAAGCCAAATATCTTGATTC	ATTACTGTGATCAAGCTTGAAACCTTTT	354
Db	129	GATAACCTCAAAACCTTTCAGAGGTTA	CAATAACAGGTGATCAAGCCGCTATTTTCT	188
QY	355	CAATCTGGGTTACCTCAACCTGTTT	TAGACAGATATGGGCACTACTGACATGATTAAT	414
Db	189	CAGTCAGGCTTGCCGGGCCCCGATT	TAGCTGAAATATGGGCCCTTATCAAGTCTGAACAA	248
QY	415	GATGGAAGATGATCAAGTGAAGTTT	TCCATAGCTATGAAACCTTATCAAACTGAAGCTA	474
Db	249	GATGGAAGATGAGACACAGCAAGATCT	CTATAGCTATGAACACTCATCAAGTTAAAGTTG	308
QY	475	CAAGGATTCAGCTACC	491	
Db	309	CAGGCCCCAACAGCTGCC	325	

Db	249	GATGGGAAGATGACACAGAAAGATTCTATAGCTATGAACCTCATCAAGTTAAGTTG	308
Qy	475	CAAGGATATCACTACC	491
Db	309	CAGGCGCAACAGCTGC	325

Search completed: July 1, 2004, 13:09:09
Job time : 365 secs

```

RESULT 15
US-09-907-969-5
? Sequence 5, Application US/09907969
? Publication No. US20030091580A1
? GENERAL INFORMATION:
? APPLICANT: Mitcham, Jennifer L.
? APPLICANT: King, Gordon E.
? APPLICANT: Algate, Paul A.
? APPLICANT: Fling, Steven P.
? APPLICANT: Retter, Marc W.
? APPLICANT: Fanger, Gary Richard
? APPLICANT: Reed, Steven S.
? APPLICANT: Vedvick, Thomas S.
? APPLICANT: Carter, Darrick
? APPLICANT: Hill, Paul
? APPLICANT: Albome, Barl
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
? OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
? FILE REFERENCE: 210121.462c8
? CURRENT APPLICATION NUMBER: US/09/907,969
? CURRENT FILING DATE: 2001-07-17
? NUMBER OF SEQ ID NOS: 596
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 5
? LENGTH: 531
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-907-969-5

```

Query Match	27.6%;	Score 137.8;	DB 10;	Length 531;
Best Local Similarity	64.7%;	Pred. No. 4.3e-34;		
Matches 205; Conservative	0;	Mismatches 112;	Indels 0;	Gaps 0;

QY 17 GGGGTCGATTCAGTCAAGTAAAGTAACAGAACCATGGCTGAGTTTCCAAACCTTTGGT 234

Db 9 GGCTGAGAGCTGCGAAGAAAGATCAGATCATGTGATGGCTAGTTTCCACAGGATGAAAT 68

QY 235 GGCAGCTGGGATATATCTGGGCCCTTAACCTGTAGAGAAAGCGAAGCATATCAGCACTTC 294

Db 69 GGAGGGCCAAATATATGGGCTATTTCATCTGAAGAGCACTAATGACATGATTAACAGTTT 128

QY 295 CATAGTTTAAAGCCAAATATCTGGATTCATTCTGGTATCAGCTCAAGTCAAAACTTTTTTTT 354

Db 129 GATTAACCTCAAACTTTCAGAGAGTATCAATTAACAGATGATCAAGCCCGTATCTTTTTCCTA 188

QY 355 CAATCTGGGTTAACCCTCAACCTGTTTTTACACAGATATGGGACACTAGTGCATGAATTAAT 414

Db 189 CAGTACAGCTCTCCGGCCCCGGTTTTAGCTGAAATATAGGCTTTATAGATCTGAACAG 248

QY 415 GATGGAAGATGATCAAGTGAAGTTTTCCAVAGCTATGAAACTTATCAACTGAAGACTA 474

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 2007.31 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_15_102

Perfect score: 459

Sequence: 1 MAITVEERAKHQFHSLKP.....KLKIKLQGYQLPSALPPVM 88

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xid
-O=/cgm2.1/USPTO.spool_p/us9720934/runat 30062004 064539 13430/app.query.fasta_1.1386
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09720934 @CGN 1 1 7509 @runat 30062004 064539 13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEDUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl :
1: gb_ba :
2: gb_hhg :
3: gb_in :
4: gb_om :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vl :
15: em_ba :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_or :
22: em_ov :
23: em_pat :
24: em_ph :
25: em_pl :
26: em_ro :
27: em_sts :
28: em_un :

29: em_vl :
30: em_hhg_hum :
31: em_hhg_inv :
32: em_hhg_other :
33: em_hhg_mus :
34: em_hhg_pin :
35: em_hhg_rod :
36: em_hhg_mam :
37: em_hhg_vrt :
38: em_sy :
39: em_hhgo_hum :
40: em_hhgo_mus :
41: em_hhgo_other :

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	459	100.0	877	6 BD124639	BD124639 Primer fo
2	459	100.0	877	6 BD126208	BD126208 Primer fo
3	459	100.0	1299	9 BC039036	BC039036 Homo sapi
4	459	100.0	1676	6 BD127640	BD127640 Primer fo
5	459	100.0	1676	9 AK075290	AK075290 Homo sapi
6	459	100.0	2079	6 BD205036	BD205036 Isolated
7	459	100.0	2131	6 AX880753	AX880753 Sequence
8	459	100.0	2131	6 BD127168	BD127168 Primer fo
9	459	100.0	2131	6 BD158570	BD158570 Primer fo
10	459	100.0	2131	9 AK027846	AK027846 Homo sapi
11	459	100.0	2131	9 AK074554	AK074554 Homo sapi
12	459	100.0	2199	9 BC058925	BC058925 Homo sapi
13	459	100.0	5195	6 BD205035	BD205035 Isolated
14	459	100.0	5199	6 BD205033	BD205033 Isolated
15	459	100.0	5287	9 AF064243	AF064243 Homo sapi
16	459	100.0	5361	9 AF114488	AF114488 Homo sapi
17	459	100.0	5458	6 BD205034	BD205034 Isolated
18	459	100.0	6439	9 AF114487	AF114487 Homo sapi
19	459	100.0	7247	9 AF064244	AF064244 Homo sapi
20	441	96.1	7493	10 AF062079	AF062079 Mus muscu
21	441	96.1	3723	10 AF132478	AF132478 Mus muscu
22	441	96.1	5145	10 AF132481	AF132481 Mus muscu
23	440	95.9	3812	10 AF132672	AF132672 Rattus no
24	440	95.9	4025	10 AF127798	AF127798 Rattus no
25	426	92.8	676	6 AX870076	AX870076 Sequence
26	426	92.8	676	6 BD150138	BD150138 Primer fo
27	409	89.1	4103	5 AF032118	AF032118 Xenopus l
28	389	84.7	481	10 AF468654	AF468654 Mus muscu
29	377	82.1	3390	9 HSM809003	HSM809003 Homo sapi
30	365	79.5	1749	9 BC020921	BC020921 Homo sapi
31	365	79.5	2085	9 BC038963	BC038963 Homo sapi
32	365	79.5	4557	9 AF182199	AF182199 Homo sapi
33	365	79.5	5828	6 BD167848	BD167848 Method fo
34	365	79.5	5828	6 AF248540	AF248540 Homo sapi
35	365	79.5	5938	9 AB033082	AB033082 Homo sapi
36	365	79.5	6091	9 AF182198	AF182198 Homo sapi
37	359	78.2	480	6 BD264675	BD264675 Compositi
38	359	78.2	480	6 AR238078	AR238078 Sequence
39	359	78.2	480	6 AR257619	AR257619 Sequence
40	359	78.2	480	6 AR283665	AR283665 Sequence
41	359	78.2	480	6 AX366293	AX366293 Sequence
42	359	78.2	531	6 BD264620	BD264620 Compositi
43	359	78.2	531	6 AR238023	AR238023 Sequence
44	359	78.2	531	6 AR257564	AR257564 Sequence
45	359	78.2	531	6 AR283610	AR283610 Sequence

RESULT 1

ALIGNMENTS

BD124639
LOCUS BD124639 877 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD124639
VERSION BD124639.1 GI:23219584
KEYWORDS JP 2002017375-A/70.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.
AUTHORS
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 70 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/70
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N5/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
source 1. 877 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.72e-53 Length: 877
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-720-934-2_COPY_15_102 (1-88) x BD124639 (1-877)
QY 1 TTPAATLTHPRVAGLUGUAGUAGALALYSHISAAPGNGINPHEHISERLEULYSFRO 20
DB 247 TGGGCGCATTAAGAGAGAAAGCGAAGCATATGACGATTCATGATTAAAGCCA 306
QY 21 ILESERGLYPHELIETHNGLYASPGINALAARGANPHEPHEGINSERGLYLEU 40
DB 307 ATATCTGATTCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 366
QY 41 GINPROVALLEUALAGINLIETTPALALEUALASPWETASANAASPGIYARGMETASP 60
DB 367 CAACTGTTTAAAGACAGATATGCGCACTAGCTGACATGAATATGATGAGAAATGAT 426
QY 61 GINVALGUPPESERILEALAMELYSEULLEYSEULYSEULNGIYTYRGINLEU 80
DB 427 CAACTGAGGTTTCCATAGCTAAGAACTTATCAAACTGAAAGCTAAGATATCAGCTA 486
QY 81 PROSERIALAENUPROVALMET 88
DB 487 CCGCTGCACTTCCCTGTCATG 510
RESULT 2
BD126208

LOCUS BD126208 877 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD126208
VERSION BD126208.1 GI:23221153
KEYWORDS JP 2002017375-A/1639.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.
AUTHORS
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 1639 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/1639
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N5/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
source 1. 877 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.72e-53 Length: 877
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-720-934-2_COPY_15_102 (1-88) x BD126208 (1-877)
QY 1 TTPAATLTHPRVAGLUGUAGUAGALALYSHISAAPGNGINPHEHISERLEULYSFRO 20
DB 247 TGGGCGCATTAAGAGAGAAAGCGAAGCATATGACGATTCATGATTAAAGCCA 306
QY 21 ILESERGLYPHELIETHNGLYASPGINALAARGANPHEPHEGINSERGLYLEU 40
DB 307 ATATCTGATTCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 366
QY 41 GINPROVALLEUALAGINLIETTPALALEUALASPWETASANAASPGIYARGMETASP 60
DB 367 CAACTGTTTAAAGACAGATATGCGCACTAGCTGACATGAATATGATGAGAAATGAT 426
QY 61 GINVALGUPPESERILEALAMELYSEULLEYSEULYSEULNGIYTYRGINLEU 80
DB 427 CAACTGAGGTTTCCATAGCTAAGAACTTATCAAACTGAAAGCTAAGATATCAGCTA 486
QY 81 PROSERIALAENUPROVALMET 88
DB 487 CCGCTGCACTTCCCTGTCATG 510
RESULT 3
BC039036
LOCUS BC039036 1299 bp mRNA linear PRI 24-DEC-2002

```

US-09-720-934-2_COPY_15_102 (1-88) x BC039036 (1-1259)

QY      1  TTPAla1leThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db      306  TGGGCCCAATCACTGATAGAGGAAAGCCAGACATGATCAGCAGTTCACATAGTTTAAAGCA 365

QY      21  IISerGlyPheIleThrGlyAspGlnAlaArgAsnGlyPheGlnSerGlyLeuPro 40
Db      366  ATATCTGATTCATTACTCGGTGATCAAGCTTAGAACTTTTCTTTTTCATCTGGGTACT 425

QY      41  GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db      426  CAACCTGTTTAGCAGCAGATATGGGCACTGACTGACATGAAATATGATGAAAGATGAT 485

QY      61  GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db      486  CAAGTGAAGTTTCCATAGCTATGATAAACTTATCAAACTGAAGCTTACAGGATATCAGCTA 545

QY      81  ProSerAlaLeuProProValMet 88
Db      546  CCTCTGCACCTTCCCCCTGTGATG 569

RESULT 4
BD127640      1676 bp      DNA      linear      PAT 18-SEP-2002
LOCUS      BD127640
DEFINITION      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      BD127640
VERSION      BD127640.1  GI:23222585
KEYWORDS      JP 2002017375-A/3071.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (Bases 1 to 1676)
AUTHORS      Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
      Makamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
      Koga,H.
      Primer for synthesizing full-length cDNA and use thereof
      Patent: JP 2002017375-A 3071 22-JAN-2002;
      HELIX RESEARCH INSTITUTE
COMMENT      OS Homo sapiens (human)
      PN JP 2002017375-A/3071
      PD 22-JAN-2002
      PF 07-JUL-2000 JP 2000253172
      PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
      PI ISHII,
      PI YURI KAWAI,AI MAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
      SHINICHI KOUTWA,
      PI TETSUJI OTSUKI,HISASHI KOGA
      PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
      PC 10",C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
      Primer for synthesizing full-length cDNA and use thereof FH Key
      Location/Qualifiers
      FT CDS
      Location/Qualifiers
      1..1676
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 5,79e-53 Length: 1676
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BD127640 (1-1676)

```

QY 1 TTPAIAIEThrValGIuGIuArgAlaYSHsAspGInGlnPheHisSerLeuYsPro 20
 Db 306 TGGGCGCATACTAGTAGAGAAAGAGCGAAGCATGATCAGAGTTCCATGTTAAAGCCA 365
 QY 21 IIESeRGIYPheIIeThrGIYAspGInAlaA-gAsnPhenPheGInSeRGIYLeuPro 40
 Db 366 AATATCGAGTTCAATTACTGATCTGATCAAGCTTGAACCTTTTTCATCAATCGGGTTACCT 425
 QY 41 GInProValIleuAlaGInIIEThrPalaleuAlaAspMetAsnAsnAspGIYArgMetAsp 60
 Db 426 CAACCGTTTTCGACAGATATGGGACACTAGCTGACATGATATGATGAGAAATGAGAT 485
 QY 61 GInValGIuPheSerIIEAlaMetIYsLeuIIEYsLeuYsLeuGInGIYTrGIuLeu 80
 Db 486 CAAGTGGAGTTTTCATGACTTGAACCTTATCAAACTTCAAACTCAAGCTACAGATATGAGCTA 545
 QY 81 ProSeRAlaIeuProProValMet 88
 Db 546 CCTCTGACACTTCCCTGTCAATG 569

RESULT 5
 AK075290
 LOCUS 1676 bp mRNA linear PRI 03-SEP-2002
 DEFINITION Homo sapiens CDNA FLJ90409 fis, clone Y79AA1000778, highly similar to Homo sapiens Intersectin long form mRNA.
 ACCESSION AK075290
 VERSION AK075290.1 GI:22761283
 KEYWORDS oligo cloning; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chiroptera; Carnivora; Primates; Catarrhini; Homidae; Homo.

1
 Isogai, T., Oca, T., Nishikawa, T., Hayashi, K., Otsuki, T., Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S., Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kojima, S., Nagahara, K., Masuno, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Niimura, Y.,
 NEDO human cDNA sequencing project

TITLE Unpublished
 JOURNAL 2 (bases 1 to 1676)
 REFERENCE Isogai, T. and Otsuki, T.
 AUTHORS Direct Submission
 JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5' - 3' end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).

FEATURES
 source location/Qualifiers

1..1676
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="Y79AA1000778"
 /cell_line="Y79"
 /cell_type="retinoblastoma"
 /clone_id="Y79AA1"
 /note="Cloning vector: pME18SFL3"

ORIGIN

Alignment Scores:
 Pred. No.: 5.79e-53
 Score: 459.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 9
 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AK075290 (1-1676)
 QY 1 TTPAIAIEThrValGIuGIuArgAlaYSHsAspGInGlnPheHisSerLeuYsPro 20
 Db 306 TGGGCGCATACTAGTAGAGAAAGAGCGAAGCATGATCAGAGTTCCATGTTAAAGCCA 365
 QY 21 IIESeRGIYPheIIeThrGIYAspGInAlaA-gAsnPhenPheGInSeRGIYLeuPro 40
 Db 366 AATATCGAGTTCAATTACTGATCTGATCAAGCTTGAACCTTTTTCATCAATCGGGTTACCT 425
 QY 41 GInProValIleuAlaGInIIEThrPalaleuAlaAspMetAsnAsnAspGIYArgMetAsp 60
 Db 426 CAACCGTTTTCGACAGATATGGGACACTAGCTGACATGATATGATGAGAAATGAGAT 485
 QY 61 GInValGIuPheSerIIEAlaMetIYsLeuIIEYsLeuYsLeuGInGIYTrGIuLeu 80
 Db 486 CAAGTGGAGTTTTCATGACTTGAACCTTATCAAACTTCAAACTCAAGCTACAGATATGAGCTA 545
 QY 81 ProSeRAlaIeuProProValMet 88
 Db 546 CCTCTGACACTTCCCTGTCAATG 569

RESULT 6
 BD205036
 LOCUS 2079 bp DNA linear PAT 17-JUL-2003
 DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
 ACCESSION BD205036
 VERSION BD205036.1 GI:33014806
 KEYWORDS JP 2002511267-A/4.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 2079)
 Korenberg, J.R. and Chen, X.N.
 Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
 Patent: JP 2002511267-A 4 16-APR-2002;
 CEDARS SINAI HEALTH SYSTEM ET AL

JOURNAL

COMMENT
 OS Homo sapiens (human)
 PN JP 2002511267-A/4
 PD 16-APR-2002
 PF 16-APR-1999 JP 2000543610
 PR 16-APR-1998 US 60/082007
 PT UJLE R KORENBERG, XIAO NING CHEN
 PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10, C12Q1/68,
 PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
 CC Isolated SH3 gene relating to myeloproliferative disorders and

CC leukemia
 CC and utilization thereof.
 FH key location/Qualifiers
 FT source 1..2079
 FT /organism="Homo sapiens (human)"
 /db_xref="taxon:9606"

FEATURES

source location/Qualifiers
 1..2079
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 7.45e-53
 Score: 459.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 6
 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BD205036 (1-2079)

QY 1 TTPAlaIleThrValGluGluArgAlaYshSAspGlnGlnPheHisSerLeuIysPro 20
Db 178 TGGGCATTAAGTCTAGAGGAAAGAGGAGCATGATCAGAGTTCATAGTTAAAGCCA 237
QY 21 IISerGlyPheIleThrGlyAspGlnAlaArgAsnPhenheGlnSerGlyLeuPro 40
Db 238 ATATCTGGATTCAATCTAGTGTGATCAAGCTAGAAACCTTTTTCATCTGGATTACCT 297
QY 41 GlnProValLeuAlaGlnIleThrPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db 298 CAACCTGTTTACACAGATATGGGCACTAGCTGACATGATATATGATGGAAGAATGGAT 357
QY 61 GlnValGluPheSerIleAlaMetIysLeuIleIysLeuIysLeuGlnGlyTyrGlnLeu 80
Db 358 CAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTGAACTACAGAGATATCAGCTA 417
QY 81 ProSerAlaLeuProProValMet 88
Db 418 CCCTCTGCACTTCCCTCTGCATG 441
RESULT 7
AX880753 2131 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 15658 from Patent EP1074617.
DEFINITION AX880753
ACCESSION AX880753 GI:40035489
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,T.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 15658 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source location/Qualifiers
1. 2131
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
347..>2131
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE92347.1"
/translation="MAQPTTPFGSLDTWALTVERAKHQQPHSLKPIGFTMGDA
KNEFFQGLDQPVLAQIWMALDMNDGMDQVPSIAMKLTKLQGLQDPSALPVM
KOQFVAISSAPFAVAPPLANGAPVIOPLAFAPHPATLTKSSFSRSGSQLNK
LOKASPDVASVPVAEMAVPOSRLKYOLFNSHDKTMSGLTGDPARTIMQSSLP
QAOIASIWNLSIDDDGKTAEFFILAMHLIDVMSGQPPVLPPIYIPSPFRVRS
GSGISVTSISVSDRLPREPVLDEQOOLEKULPTEDEKRENERGNELEERQA
LLEQORKEQLAQLERAEQERKERERQERKQLELEKQLEKQRELBQERERK
EIERREAKRELERQLEWERNRQELNQRNEQEDIVLAKKTKLEBELDAND
KHLQLEGLQDLCRIITTORQELSTNSRSLRAETLTHLQOOLQESQOMIGRLIPK
QIINDOLKOVQNSIHRSDSVTLKRALFAKRLAQHLRDQIDEVEKETRSKLQEDIDF
NNQLEKEIREINHQQLQKQKMEARLQK"
ORIGIN
Alignment Scores:
Pred. No.: 7.67e-53 Length: 2131
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 6
US-09-720-934-2_copy_15_102 (1-88) x AX880753 (1-2131)
QY 1 TTPAlaIleThrValGluGluArgAlaYshSAspGlnGlnPheHisSerLeuIysPro 20
Db 178 TGGGCATTAAGTCTAGAGGAAAGAGGAGCATGATCAGAGTTCATAGTTAAAGCCA 237

Db 389 TGGGCATTAAGTCTAGAGGAAAGAGGAGCATGATCAGAGTTCATAGTTAAAGCCA 448
QY 21 IISerGlyPheIleThrGlyAspGlnAlaArgAsnPhenheGlnSerGlyLeuPro 40
Db 449 ATATCTGGATTCAATCTAGTGTGATCAAGCTAGAAACCTTTTTCATCTGGATTACCT 508
QY 41 GlnProValLeuAlaGlnIleThrPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db 509 CAACCTGTTTACACAGATATGGGCACTAGCTGACATGATATATGATGGAAGAATGGAT 568
QY 61 GlnValGluPheSerIleAlaMetIysLeuIleIysLeuIysLeuGlnGlyTyrGlnLeu 80
Db 569 CAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTGAACTACAGAGATATCAGCTA 628
QY 81 ProSerAlaLeuProProValMet 88
Db 629 CCCTCTGCACTTCCCTCTGCATG 652
RESULT 8
BD127168 2131 bp DNA linear PAT 18-SEP-2002
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD127168
ACCESSION BD127168 GI:23222113
VERSION JP 2002017375-A/2599.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2131)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2599 22-JAN-2002;
HELIIX RESEARCH INSTITUTE
OS Homo sapiens (human)
COMMENT PN JP 2002017375-A/2599
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,
PC C12P21/02,C12O1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
source location/Qualifiers
FT CDS (205).. (2130).
1. 2131
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7.67e-53 Length: 2131
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 6
US-09-720-934-2_copy_15_102 (1-88) x BD127168 (1-2131)
QY 1 TTPAlaIleThrValGluGluArgAlaYshSAspGlnGlnPheHisSerLeuIysPro 20
Db 247 TGGGCATTAAGTCTAGAGGAAAGAGGAGCATGATCAGAGTTCATAGTTAAAGCCA 306

```

QY 115SerGlyPheIleThrGlyAspGlnAlaArgAsnPhenheGlnSerGlyLeuPro 40
DB 307 AATATCGATTCACTACTGATGATCAAGCTAGAACTTTTTCATCTGGGTTACT 366
QY 41 GlnProValLeuAlaGlnIleTrrPalalaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 367 CAACCTGTTTTCACAGATATGGGACATGAGCTGACATGATATATATGAAAGAAATGGAT 426
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 427 CAAGTGAGCTTTCCATAGCTATGAACCTATCAAACTGACAACTGACAGATATTCAGCTA 486
QY 81 ProSerAlaLeuProProValMet 88
DB 487 CCCTCGACCTCCCTGTCATG 510

RESULT 9
LOCUS BD158570 2131 bp DNA linear PAT 17-JUN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD158570
VERSION BD158570.1 GI:27864328
KEYWORDS JP 2002191363-A/13413.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 2131)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 13413 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/13413
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI MAKAMATSU,
PI KEIICHI NAGAI, TETSUKI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
Source Location/Qualifiers
1..2131
/organism="Homo sapiens"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7.67e-53 Length: 2131
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BD158570 (1-2131)
QY 1 TrrPalalThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 389 TGGGCCCTAATCTGAGGAAAGAGCCAGCATGATGACAGCTTCCATAGTTTAAAGCCA 448
QY 21 Il5SerGlyPheIleThrGlyAspGlnAlaArgAsnPhenheGlnSerGlyLeuPro 40
DB 449 AATATCGATTCACTACTGATGATCAAGCTAGAACTTTTTCATCTGGGTTACT 508

```

```

QY 41 GlnProValLeuAlaGlnIleTrrPalalaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 509 CAACCTGTTTTCACAGATATGGGACATGAGCTGACATGATATATATGAAAGAAATGGAT 568
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 569 CAAGTGAGCTTTCCATAGCTATGAACCTATCAAACTGACAACTGACAGATATTCAGCTA 628
QY 81 ProSerAlaLeuProProValMet 88
DB 629 CCCTCGACCTCCCTGTCATG 652

RESULT 10
LOCUS AK027846 2131 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ14940 fis, clone PLACE1010942, highly similar
to Homo sapiens intersecin long isoform (ITSN) mRNA.
ACCESSION AK027846
VERSION AK027846.1 GI:14042823
KEYWORDS oligo cloning: fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K.,
Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Ishii,S., Kawai,Y.,
Saito,K., Yamamoto,J., Makamatsu,A., Nakamura,Y., Nagahari,K.,
Masuko,Y. and Kanehori,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2131)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-1975, Fax:81-438-52-3586)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction,
5'- & 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Source Location/Qualifiers
1..2131
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1010942"
/tissue_type="placenta"
/clone_id="PLACE1"
/note="Cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.: 7.67e-53 Length: 2131
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AK027846 (1-2131)
QY 1 TrrPalalThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 389 TGGGCCCTAATCTGAGGAAAGAGCCAGCATGATGACAGCTTCCATAGTTTAAAGCCA 448
QY 21 Il5SerGlyPheIleThrGlyAspGlnAlaArgAsnPhenheGlnSerGlyLeuPro 40
DB 449 AATATCGATTCACTACTGATGATCAAGCTAGAACTTTTTCATCTGGGTTACT 508

```

QY 41 GlnProValIleuAlaGlnIleTyrPalAlaValAlaAspMetCAsnAsnAspGlyArgMetAsp 60
 DB 509 CAACCTGTTTACACAGATATGCGACAGCTGACATGATATGATGAGAAAGATGAT 568
 QY 61 GlnValGlnPheSerIleAlaMetCysLeuIleIleValLeuIleValGlnGlyTyrGlnIleu 80
 DB 569 CAAGTGGAGTTTCCATGCTATGAACTTATCAAACTGAAGCTACAGGATATACAGCTA 628
 QY 81 ProSerAlaIleuProProValMet 88
 DB 629 CCTCTGCACCTTCCCTCTGCATG 652

RESULT 11
 AK074554 2131 bp mRNA linear PRI 03-SEP-2002
 AK074554
 LOCUS Homo sapiens cDNA FLJ90073 fis, clone HEMBA1004110, highly similar
 DEFINITION to Homo sapiens interectin short form mRNA.
 ACCESSION AK074554 GI:22760070
 VERSION AK074554.1
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS Isegai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T., Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S., Kawai-Hito, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kojima, S., Nagahara, K., Masuno, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsuma, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Ninomiya, K.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2131)
 AUTHORS Isegai, T. and Otsuki, T.
 JOURNAL Direct Submission
 COMMENT Submitted (25-MAR-2002) Takao Isegai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

FEATURES
 SOURCE 1..2131
 Location/Qualifiers
 1..2131
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEMBA1004110"
 /issue_type="whole embryo, mainly head"
 /clone_lib="HEMBA1"
 /dev_stage="embryo, 10 weeks"
 /note="Cloning vector: PME18SFL3"

ORIGIN
 Alignment Scores: Length: 2131
 Pred. No.: 7.67e-53
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 9

US-09-720-934-2_COPY_15_102 (1-88) x AK074554 (1-2131)

QY 1 TrrAlaIleThrValGlnGluArgAlaLysHisAspGlnPheHisSerIleuLysPro 20
 DB 247 TGGGCGATTAAGCTGAGAGAAAGAGGAGCATGATCAGACGTTCCATGTTTAAAGCA 306

QY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
 DB 307 ATATCTGAGTTTCAATTAAGTATGATGATCAGCTAGAACTTTTTCATCTGGGTTACCT 366
 QY 41 GlnProValIleuAlaGlnIleTyrPalAlaValAlaAspMetCAsnAsnAspGlyArgMetAsp 60
 DB 367 CAACCTGTTTACACAGATATGCGACAGCTGACATGATATGATGAGAAAGATGAT 426
 QY 61 GlnValGlnPheSerIleAlaMetCysLeuIleIleValLeuIleValGlnGlyTyrGlnIleu 80
 DB 427 CAAGTGGAGTTTCCATGCTATGAACTTATCAAACTGAAGCTACAGGATATACAGCTA 486
 QY 81 ProSerAlaIleuProProValMet 88
 DB 487 CCTCTGCACCTTCCCTCTGCATG 510

RESULT 12
 BC058925 2199 bp mRNA linear PRI 05-NOV-2003
 BC058925
 LOCUS Homo sapiens interectin 1 (SH3 domain protein), mRNA (cDNA clone
 DEFINITION IMAGE:443129), partial cds.
 ACCESSION BC058925 GI:37589134
 VERSION BC058925
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS Klausner, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G., Strausberg, R.L., Sanger, C., Wagner, J., Shenmen, C.M., Schuler, G.D., Klausner, R.D., Collins, F.S., Bonaldo, M.F., Casavant, T., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T., Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwen, P.J., McKernan, K.D., Malek, U.A., Gunaratne, P.H., Richards, S., Worrley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Bulky, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 2199)
 AUTHORS Strausberg, R.
 JOURNAL Direct Submission
 COMMENT Submitted (01-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Offices, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.mci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guo, Letitia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Sze Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nese, Pawan Pandoh, Anna-Ilisa Prabh, Parvaneh Saeedi, Jacqueline

Schein, Duane Smalrus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/HLN at: <http://image.llnl.gov>
Series: IRAC Plate: 119 Row: B Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796.

FEATURES

source

1. 2199
Location/Qualifiers

gene

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:443129"

/tissue_type="liver, adenocarcinoma"

/clone_id="N1H MCC_90"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. >2199

/gene="ITSN1"

/note="synonym: ITSN, SH3DLA, SH3P17"

/db_xref="locusid:6453"

/db_xref="MIM:602442"

296. .>2199

/gene="ITSN1"

/codon_start=1

/product="ITSN1 protein"

/protein_id="AAH58925.1"

/db_xref="GI:37589135"

/db_xref="MIM:602442"

/db_xref="locusid:6453"

/db_xref="MIM:602442"

/translation="NAOPEPTPGSLDIWAIIVTERAKHDQFHSIKPISGRTGDOA
RNFPSQSLPQVLQAIWALAMNNNDMDVSEIIMKLIKIKIGYOLPALPDM
KOOPVISAAPFGMGTASMPPLTAVA PVPMSGTIVSVPTAVPEPLAN
GAPVIOPLPAFNAHATLPKSSSSRSQSLTKIQKASIVASVPAEVAVP
OSSRLTKOLENSHDKTWSHITGPQARTILMOSSLPQAQLASIVNLSDIDQDKLTA
BEFIIAMKIDIVAMGQPLPVPYLPPEYIPSPFRVSGSGISVISTVDRLPEPV
LEDEQOQLEKLPVFEDEKRENERGNILEKROALCOORKEOEALORLEPEV
RKERROEROEROLEKOLEKORELROREERKEIERRAKRLEPELOEME
RNROELNORKEGEDIIVLKAKKTELEFLEALNDKQLEBGLQDTRCRLTTO
EIESNKREIRIAHITLQOOLQSQMLGRLPERKQILNDQLQVONSISHSSTV
TLKRLAKELARQHLRDQDLEVEKERSKLQEIIDFNOLKELEIHNKOLOKOKS
MEARLKRRKK"

335. .619

/gene="ITSN1"

/note="EH: Region: Eps15 homology domain"

/db_xref="CDD:smart0027"

935. .1171

/gene="ITSN1"

/note="EH: Region: Eps15 homology domain"

/db_xref="CDD:smart0027"

ORIGIN

Alignment Scores:

Pred. No.: 7.96e-53 Length: 2199
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BC058925 (1-2199)

QY 1 TTPALAIETHRVAGLUGUAGALALYSHISAPGNGINPHEHISSEULEUYPRO 20
DB TGGCCCACTAATCTGAGAGAAAGAGCATATGACAGTTCAGTTTAAAGCA 397
QY 21 ILESERGLYPHEIETHRVAGLUGUAGALALYSHISAPGNGINPHEHISSEULEUYPRO 40
DB ATATCTGATTCATTAATCTGAGAGAAAGAGCATATGACAGTTCAGTTTAAAGCA 457
QY 41 GINPROVALLEUALAGINIIETHRVAGLUGUAGALALYSHISAPGNGINPHEHISSEULEUYPRO 60

Db

458 CAACCTGTTTACCAAGATATGGCACTGACATGATATATATGAGAAATGAT 517

QY

61 GINVALGLUPHESERILEALAMELYSLEULELYSEULENGINGLYTGGINLEU 80

Db

518 CAAGTGAAGTTTCCATGACTATGAAACTATCAAACTGAGCTACAGATATGACTA 577

QY

81 PROSERALALEUPROPROVALMET 88

Db

578 CCCTGTGACACTTCCCTGTGACTG 601

RESULT 13

BD205035

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

PD

PF

PI

PC

PC

CC

CC

CC

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

ORIGIN

Alignment Scores:

Pred. No.: 2.18e-52 Length: 5195
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BD205035 (1-5195)

QY 1 TTPALAIETHRVAGLUGUAGALALYSHISAPGNGINPHEHISSEULEUYPRO 20
DB TGGCCCACTAATCTGAGAGAAAGAGCATATGACAGTTCAGTTTAAAGCA 340
QY 21 ILESERGLYPHEIETHRVAGLUGUAGALALYSHISAPGNGINPHEHISSEULEUYPRO 40
DB ATATCTGATTCATTAATCTGAGAGAAAGAGCATATGACAGTTCAGTTTAAAGCA 400
QY 41 GINPROVALLEUALAGINIIETHRVAGLUGUAGALALYSHISAPGNGINPHEHISSEULEUYPRO 60
DB CAACCTGTTTACCAAGATATGGCACTGACATGATATATATGAGAAATGAT 460

Qy	61	GINValGlnPheSerIleAlaMetIleValLeuGlnIleValLeuGlnGlyTyrGlnLeu	80
Db	461	CAAGTGGAGGTTTCCATAGCTATAGAACTTATCAACTGAGCTACAGATACAGCTA	520
Qy	81	ProSeRalLeuProProValMet	88
Db	521	CCCTCTGCACCTTCCCTGTCATG	544
RESULT 14			
BD205033		5199 bp	DNA
BD205033			linear
LOCUS			PAT 17-JUN-2003
DEFINITION			Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
ACCESSION			BD205033
VERSION			BD205033.1
KEYWORDS			GI:33014803
SOURCE			JP 2002511267-A/1.
ORGANISM			Homo sapiens (human)
REFERENCE			Homo sapiens
AUTHORS			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE			1 (bases 1 to 5199)
JOURNAL			Korenberg,J.R. and Chen,X.N.
COMMENT			Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
			Patent: JP 2002511267-A 1 16-APR-2002;
			CEDARS. SIVAI HEALTH SYSTEM ET AL
			OS Homo sapiens (human)
			PN JP 2002511267-A/1
			PD 16-APR-2002
			PF 16-APR-1999 JP 2000543610
			PR 16-APR-1998 US 60/082007
			PI JULIE R. KORENBERG, XIAO NING CHEN
			PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
			PC C12Q1/68
			PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
			CC Isolated SH3 gene relating to myeloproliferative disorders and
FEATURES			
source			
ORIGIN			
Alignment Scores:			
Pred. No.:	2,18e-52	Length:	5199
Score:	459.00	Matches:	88
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-720-934-2_COPY_15_102 (1-88) x BD205033 (1-5199)			
Qy	1	TPPAlaIleThrValGlnGlnArgAlaIleValSH3ASpGlnGlnPheHisSerLeuIleValPro	20
Db	250	TGGGCGCACTAAGTGTAGAGGAAAGAGCGAAGCATGATGACAGTTCATGTTTAAAGCCA	309
Qy	21	IIeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro	40
Db	310	AATATCGAGTCAATTACTGCTGATGACGCTAGAACTTTTTCATATCTGGGTTACCT	369
Qy	41	GlnProValIleValGlnIleThrPalaIleValAspMetAsnAsnAspGlyArgMetAsp	60
Db	370	CAACCGTTTTCGACACGATATGGGACCATGATGACATGAATATATGGAAGAATGAT	429
Qy	61	GlnValGlnPheSerIleAlaMetIleValLeuGlnIleValLeuGlnGlyTyrGlnLeu	80

Db	Cy	Db	RESULT 15	LOCUS	AF064243	5287 bp	mRNA	linear	PRI 21-NOV-1998				
430	ProSerAlaLeuProProValMet 86	CCCTGAGAGTTTCCATAGCTAAGAACTTATCAACTAAGCTACAAAGATATCAGCTA	489	AF064243	5287 bp	mRNA	linear	PRI 21-NOV-1998					
				AF064243	Homo sapiens interseclin short form mRNA, complete cds.								
	490	CCCTGCACTTCCCTCCCTGATCAG	513	AF064243.1	GI:3859852								
					Homo sapiens (human)								
					Homo sapiens								
					Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
					Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
					1 (bases 1 to 5287)								
					Guipponi M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and Antonarakis, S.B.								
					Two isoforms of a human interseclin (ITSN) protein are produced by brain-specific alternative splicing in a stop codon								
					Genomics 53 (3), 369-376 (1998)								
					9799604								
					2 (bases 1 to 5287)								
					Guipponi M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and Antonarakis, S.B.								
					Direct Submission								
					Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue								
					Michel-Sever, Geneva 4 CH-1211, Switzerland								
					Location/Qualifiers								
					1..5287								
					/organism="Homo sapiens"								
					/mol_type="mRNA"								
					/db_xref="taxon:9606"								
					/chromosome="21"								
					/map="21q22.1-q22.2"								
					/tissue_type="brain"								
					/dev_stage="fetus"								
					107..3769								
					/codon_start=1								
					/product="interseclin short form"								
					/protein_id="AAC78610.1"								
					/db_xref="GI:3859853"								
					/translation="MAQPTPEFGSLDIWATYBERAKHDQPSLKPISGFTGDDQARNFPGSLQPQVLAQIMALADNNDNRMOQVERSIAMKLIKLGQYLPQSLAPVPMKQCVLAISSAPPFGMGISLAWPPLTVAAPVPMGSIPLVAGMSPTLVSSVPTAAVPLMAGAPVIOPLPAFHAPRAVTLPEKSSFSFGSGSLQNTPLQKAGSDVAVSPVEMAVPQSSPLKPEKPLNSHDKTMSGLTGSGRTITMOSLIPQALASITNLSIDIDQDKLTAYCEFLIAHLITLVNMSGCLPVPVLPPIIPSPFRVRSQSGISVTSISVDORLPBEPVLEEDQQQLERKLPTVFEDRKRENTENGLELEKRRQALLBQRKEQERLQAULEABEDNRKREERQERKQLELEKQERLEKRELEKREERREIERERAKLEKROULEMQRNRREOLNQRNKEQEDIVLVKAKKTLPEFLLELNDKQLEBKLODIRLTQFRQELIESNKRERELRIAEITHLQOOLQESQOMGLRIPEKQIINDQLQVQMSLHSDSLVTLKRALEKELAROHIRNDQLEDEKERTSKLOEIDINNOLEKEREIHNKQOOLQKSKSMEBERLKQEKQERKLTLEKQEKERAGRRARQQRDMQLEHNOQDEHORPRKLEHEKTLKREESVRRKQDEERKQEAQDKLGLPHQERPAKPAVQAPMSTAKGEPLTISQENVKVYVFRALYPPESSHDEITIQPGDIVVWKGWDESGTGEPMUGGLKQKGTQWFPFRNVAEKIPENEVDPAVKEPTVTDSTSAAPKLALRETPALVSSBPTTPNNWADFSSTWPTNTEKPEKENDNMDAMAQPSLTVPSAGQLRQSRAPATPATSFSPPVLGQGEKVBVGLQDAQALYIPWRKAKDNHLENKKNVITVILEQODMMWGEVQGGKGFPSVKVLSIGBTRKTSMDNSGSESPASLKRVAAPRAKRVYSGEERLAMYTESSEODLTPOQGDVILILYTKKQDWWTGTGVGDKAGVPPSNVRLKDSGSGTAGTGLGKKPELAAQVLAAYTATGEEQQLTALPAQQLILIRKNPGWMEGELQAGKRRQIGKPPANVYKVLNPSTISVTPPEPKETLAAVACOVIGMAYDTAQNDELEAFNKGQIINVLINEKDEPDWKGENVNGLEPSNYXKLTITMDPSQQ"								
					misc_feature	167..406							
					/note="encodes EH domain"								
					misc_feature	767..1936							
					/note="encodes EH domain"								
					misc_feature	2324..2524							

misc_feature /note="encodes SH3 domain"
2843..3019
/note="encodes SH3 domain"
misc_feature 3110..3286
/note="encodes SH3 domain"
misc_feature 3326..3520
/note="encodes SH3 domain"
misc_feature 3569..3748
/note="encodes SH3 domain"
ORIGIN

Alignment Scores:
Pred. No.: 2.22e-52 Length: 5287
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * AF064243 (1-5287)

QY 1 TTPALAIETHRVALGLUGLVRGALALYSHISASPGINGLNPHETISERILEUYSPro 20
DB 149 TGGGCCATTAAGAGGAAAGAGGAGCATGATCAGCAGTCCATAGTTAAAGCCA 208
QY 21 ILESERGLYPHEIETHRGLYASPGINALARGASNPHEPHEGLNSERGLYLEUPro 40
DB 209 ATATCTGGATTCATTACTGTGATCAAGCTAGAACTTTTTCATCTGGTTACCT 268
QY 41 GINPROVALLEUALGINLETTPALALEUALASPMETASNASNASPGIYARGMETASP 60
DB 269 CAACCTGTTTTCAGACAGATATGGGCACACTGACATGATAATATATGAGAGATGGAT 328
QY 61 GINVALGLUPHESERILEALMETLYSLEUILLEYSLEUYSLGNGLYTYRGINLEU 80
DB 329 CAAGTGGAGTTTCCATTAGCTTGAAGACTTATCAAACTGAAAGCTACAAAGATATCAGCTA 388
QY 81 PROSERALALEUProPROVALMET 88
DB 389 CCTCTGCACCTTCCCCCTGTGATG 412

Search completed: July 1, 2004, 16:14:04
Job time : 2015.31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 2647.21 Seconds
(without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_15_102
Perfect score: 459
Sequence: 1 WAITEERAKHQDFHSLKP.....KLKIKLQGYQLPSALPPYM 88

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlp
-DB=CGM2_1/USPTO.spool.p/US09700934/runat.30062004.064540.13442/app.query.fasta_1.1386
-DB-RST -QEXT=fastcap -SUFFIX=ret -MINMATCH=0.1 -LOOPEC=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09700934 @CGM 1 1 12421 @runat.30062004.064540.13442 -NCPU=6 -ICPU=3
-NO MMAP -LARGEJOURNEY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rdd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	459	100.0	477	14	CB269094
2	459	100.0	544	9	AL711737
3	459	100.0	547	10	BF194430
4	459	100.0	734	14	CB269754
5	459	100.0	896	13	BQ941336
6	459	100.0	979	13	BQ941411
7	459	100.0	1089	12	BQ050397
8	459	100.0	2079	11	BC013578
9	459	100.0	2126	11	BC020269
10	451	98.3	523	13	B0381021
11	451	98.3	907	13	B0128229
12	441	96.1	445	9	AI852079
13	441	96.1	475	10	BE199657
14	441	96.1	488	9	AI481705
15	441	96.1	501	9	AI852070
16	441	96.1	510	10	BE852536
17	441	96.1	533	29	CG579249
18	441	96.1	600	14	CA528645
19	441	96.1	662	14	CF536420
20	441	96.1	698	14	CB519152
21	441	96.1	728	13	BY731854
22	441	96.1	761	14	CF729634
23	441	96.1	8385	11	BC062938
24	440	95.9	411	9	AI549192
25	440	95.9	811	12	BG829540
26	432	94.1	750	12	BG118422
27	431	93.9	643	13	BK470886
28	431	93.9	766	14	CF743758
29	426	92.8	723	10	BE786696
30	423	92.2	624	13	BY713771
31	420	91.5	359	9	AI853008
32	410	89.3	882	9	AL792490
33	410	89.3	892	9	AL870708
34	410	89.3	534	12	BG160235
35	409	89.1	550	12	B0623100
36	409	89.1	400	9	AL964682
37	365	79.5	422	14	CA388981
38	365	79.5	699	14	CA391925
39	365	79.5	399	9	AA217338
40	359	78.2	404	9	AI390576
41	359	78.2	490	9	AI646131
42	359	78.2	556	12	BG144398
43	359	78.2	565	9	AI809238
44	359	78.2	619	10	BB614005
45	359	78.2			

ALIGNMENTS

RESULT 1
LOCUS CB269094 477 bp mRNA linear EST 20-FEB-2003
DEFINITION 1008001 Human Fat Cell 5'-Stretch plus cDNA library Homo sapiens
ACCESSION CB269094
VERSION CB269094.1 GI:28443679
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 477)

AUTHORS Yang, R.-Z., Shuldiner, A. and Gong, D.-W.
 TITLE EST analysis of human adipose gene expression
 JOURNAL Unpublished (2002)
 COMMENT Contact: Gong Da-Wei
 Division of Endocrinology, Diabetes and Nutrition
 University of Maryland
 660 Redwood St., HH497, Baltimore, MD 21201, USA
 Tel: 410 706 1672
 Fax: 410 706 1622
 Email: dgong@medicine.umaryland.edu

PCR Primers
 FORWARD: CTCGGGAAGCGCGCATTGTGTGT
 BACKWARD: AATACGACTCATTAGGCGCAATTGG
 Seq primer: GTTGGTACCCGGAAATTC.
 Location/Qualifiers
 1..477
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /sex="Male and Female"
 /tissue_type="Adipose"
 /clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
 /note="Vector: lambdaTriplex"

ORIGIN

Alignment Scores:
 Pred. No.: 1,678-53 Length: 477
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x CB269094 (1-477)

QY 1 TTPALAIETNRVALGUGUARGALALYSHISASPGINGINPHEHISERLEUYSRPRO 20
 DB 4 TGGGCATTAACCTGTAAGAGAAAGAGCGAAGCATGATCGAGTCCATAGTTAAAGCA 63
 QY 21 ILESERGLYPHELIETHRGLYASPGINALARGANPHEPHEGINSERGLYLEUPRO 40
 DB 64 ATATCTGATTCATTAATCTGATCAAGTAAAGTAACTTTTTCATCTGGGTTACCT 123
 QY 41 GINPROVALLEUALAGNIIETIPALALEUALAASPMEASASAPGIVARGMECASP 60
 DB 124 CAACCTGTTTAGACACAGATATGGCACTAGCTGACATGATTAATGATGAAGATGAT 183
 QY 61 GINVALGIUPHESERTLEALAMELISLEULEYLSLEULYSLGUNGILYTYIGINLEU 80
 DB 184 CAAGTGGAGTTTCCATGACTATGAACTATCAAACTGAAGCTACAAAGATATCAGCTA 243
 QY 81 PROSERIALAENUPROPROVALMET 88
 DB 244 CCTCTGCACCTTCCCCCTGTCATG 267

RESULT 2
 LOCUS AL711737 544 bp mRNA linear EST 04-SEP-2003
 DEFINITION DKFZP686K1884_r1_686 (synonym: hicc3) Homo sapiens cDNA clone
 ACCESSION DKFZP686K1884_5, mRNA sequence.
 VERSION AL711737
 KEYWORDS AL711737.1 GI:19695092
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCES
 1 (bases 1 to 544)
 Duesterhoeft, A., Lauber, J., Mewes, H.W., Gaassenhuber, J. and
 Wiemann, S.
 EST (Duesterhoeft, et al.)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: MIPS

ORIGIN

Alignment Scores:
 Pred. No.: 1,986-53 Length: 544
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AL711737 (1-544)

QY 1 TTPALAIETNRVALGUGUARGALALYSHISASPGINGINPHEHISERLEUYSRPRO 20
 DB 216 TGGGCATTAACCTGTAAGAGAAAGAGCGAAGCATGATCGAGTCCATAGTTAAAGCA 275
 QY 21 ILESERGLYPHELIETHRGLYASPGINALARGANPHEPHEGINSERGLYLEUPRO 40
 DB 276 ATATCTGATTCATTAATCTGATCAAGTAAAGTAACTTTTTCATCTGGGTTACCT 335
 QY 41 GINPROVALLEUALAGNIIETIPALALEUALAASPMEASASAPGIVARGMECASP 60
 DB 336 CAACCTGTTTAGACACAGATATGGCACTAGCTGACATGATTAATGATGAAGATGAT 395
 QY 61 GINVALGIUPHESERTLEALAMELISLEULEYLSLEULYSLGUNGILYTYIGINLEU 80
 DB 396 CAAGTGGAGTTTCCATGACTATGAACTTATCAAACTGAAGCTACAAAGATATCAGCTA 455
 QY 81 PROSERIALAENUPROPROVALMET 88
 DB 456 CCTCTGCACCTTCCCCCTGTCATG 479

RESULT 3
 LOCUS BF194430 547 bp mRNA linear EST 02-NOV-2000
 DEFINITION 246586 MARS 2P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BF194430
 VERSION BF194430.1 GI:11077799
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa

ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 REFERENCES
 1 (bases 1 to 547)
 Fahrnenkung, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
 Vallet, J., Wise, T., Rohrer, G.A., Pettes, G., Sultana, R.,
 Quackenbush, J. and Keele, J.W.
 Porcine gene discovery by normalized cDNA-library sequencing and
 EST cluster assembly
 Mamm. Genome 13 (8), 475-478 (2002)

TITLE EST cluster assembly
 JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
 MEDLINE 22213789

PubMed 12226715
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCGCCATCAGCAGC
Plate: 76 Row: E Column: 21
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
1. 547
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2Pig"
/note="Vector: PCMV SPORT6; Site 1: Not; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

ORIGIN

Alignment Scores:
Pred. No.: 1,99e-53 Length: 547
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BF194430 (1-547)

QY 1 TTPAlalleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db 185 TGGGCCATAACTGATAGAGAAAGCAAGCATGATCAGCAGTTCCATAGCTTAAAGCCA 244
QY 21 TLeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db 245 ATATCTGATTTACTTACTGATGATCAAGTAAGAACTTTTTCATCTGGGTAACTT 304
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
Db 305 CAACCTGTTTTCAGACAGATATGGCGCTGCTGATGATATATGACGGAAGATGGAT 364
QY 61 GlnValGlnPheSerIleAlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 365 CAAGTGGAGTTTTCATAGCATGAACTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTC 424
QY 81 ProSerAlaLeuProProValMet 88
Db 425 CCTCTCGCACTTCCCTCTGTATG 448

RESULT 4
LOCUS CB269754 734 bp mRNA linear EST 20-FEB-2003
DEFINITION 1008661 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
ACCESSION CB269754
VERSION CB269754.1 GI:28444339
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 734)
AUTHORS Yang, R.-Z., Shuldiner, A. and Gong, D.-W.
TITLE EST analysis of human adipose gene expression

JOURNAL Unpublished (2002)
COMMENT Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR Primers
FORWARD: CTCGGAGAGCGCCCATTTGTGTGT
BACKWARD: AATACGACTACTATGAGCGCATTTGG
Seq primer: GTTGATACCCGGAATTC.
Location/Qualifiers
1. 734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/note="Vector: lambda triplex"

ORIGIN

Alignment Scores:
Pred. No.: 2,91e-53 Length: 734
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x CB269754 (1-734)

QY 1 TTPAlalleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db 163 TGGGCCATAACTGATAGAGAAAGCAAGCATGATCAGCAGTTCCATAGCTTAAAGCCA 222
QY 21 TLeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db 223 ATATCTGATTTACTTACTGATGATCAAGTAAGAACTTTTTCATCTGGGTAACTT 282
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
Db 283 CAACCTGTTTTCAGACAGATATGGCGCTGCTGATGATATATGACGGAAGATGGAT 342
QY 61 GlnValGlnPheSerIleAlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 343 CAAGTGGAGTTTTCATAGCATGAACTTATCAAGCTCAAGCTCAAGCTCAAGCTCA 402
QY 81 ProSerAlaLeuProProValMet 88
Db 403 CCTCTCGCACTTCCCTCTGTATG 426

RESULT 5
LOCUS BQ941336 896 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8741326 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6420600
ACCESSION BQ941336
VERSION BQ941336.1 GI:22356814
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2594 row: k column: 01
High quality sequence stop: 763.

FEATURES

source

1. 896

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:420600"

/tissue_type="large cell carcinoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_18"

/note="Organ: lung; Vector: pOT7; Site_1: XhoI; Site_2:

EcoRI; CDNA made by oligo-dr priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GCGACGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 3,766-53 Length: 896
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BQ941336 (1-896)

```

QY 1 TTPAIAIETHRVAGLUGUARGALALYSHASPGINGINPHEHSERLEUYSPro 20
DB 252 TGGGCCATTAACCTGTAAGAGAAAGCGAAGCATGATCGACATTCATGTTAAAGCCA 311
QY 21 ILESERGIYPHEIETHRGLYASPGINAAARGAENPHEPHEGINSERGIYLEUPro 40
DB 312 ATATCTGATTCATTAACCTGTAAGAGAAAGCATGATTCATGTTAAAGCCA 371
QY 41 GINPROVALLEUAAGINLETTPALALEUAIAASPMETASNAASPGIYARMEtasp 60
DB 372 CAACCTGTTTAGACACATATGCGCACTAGCTGACATGATTAAGATGAAGATTGAT 431
QY 61 GINVALGIUPHESERILEALMETLYSLEUILEYSLEUYSLEUGINGLYTYRGINLeu 80
DB 432 CAAGTGGAGTTTTCATAGCTATGAAACTTATCAAACTGAAAGCTACAAAGATATCAGCTA 491
QY 81 PROSERALALEUPROPROVALMET 88
DB 492 CCTCTGCACTTCCCTGTCATG 515

```

RESULT 6
BQ941411
LOCUS BQ941411 979 bp mRNA linear EST 21-AUG-2002
DEFINITION IMAGE:6205905 5', mRNA sequence.
ACCESSION BQ941411
VERSION BQ941411.1 GI:22356889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Molecular: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 979)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

FEATURES

source

1. 979

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6205905"

/sex="male"

/tissue_type="sciatic nerve"

/dev_stage="adult, 70 yr"

/lab_host="DH10B"

/clone_lib="Lupski sciatic nerve"

/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:

NotI; Site_2: SalI; CDNA made by oligo-dr priming.

Directionally cloned using the following adaptors:

5'-GACTAGTCTTACATGCGAGCGCGCCCT(15)-3' and

5'-TCGACCCACGCGTCCG-3'

1 kb for average insert length 1.87 kb. This is a primary

library, non-amplified. Library constructed by Life

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor

College of Medicine) and is available through Life

Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 4,226-53 Length: 979
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BQ941411 (1-979)

```

QY 1 TTPAIAIETHRVAGLUGUARGALALYSHASPGINGINPHEHSERLEUYSPro 20
DB 328 TGGGCCATTAACCTGTAAGAGAAAGCGAAGCATGATCGACATTCATGTTAAAGCCA 387
QY 21 ILESERGIYPHEIETHRGLYASPGINAAARGAENPHEPHEGINSERGIYLEUPro 40
DB 388 ATATCTGATTCATTAACCTGTAAGAGAAAGCATGATTCATGTTAAAGCCA 447
QY 41 GINPROVALLEUAAGINLETTPALALEUAIAASPMETASNAASPGIYARMEtasp 60
DB 448 CAACCTGTTTAGACACATATGCGCACTAGCTGACATGATTAAGATGAAGATTGAT 507
QY 61 GINVALGIUPHESERILEALMETLYSLEUILEYSLEUYSLEUGINGLYTYRGINLeu 80
DB 508 CAAGTGGAGTTTTCATAGCTATGAAACTTATCAAACTGAAAGCTACAAAGATATCAGCTA 567
QY 81 PROSERALALEUPROPROVALMET 88
DB 568 CCTCTGCACTTCCCTGTCATG 591

```

RESULT 7
BQ050397
LOCUS BQ050397 1089 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT 7050803 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5784342
5', mRNA sequence.
ACCESSION BQ050397
VERSION BQ050397.1 GI:19809737
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 1089)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://imgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabds-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM12869 row: 1 column: 07
High quality sequence stop: 612.
Location/Qualifiers

FEATURES

1..1089
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5784342"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6, Site 1: NotI, Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."

ORIGIN

Alignment Scores:
Pred. No.: 4.84e-53 Length: 1089
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BQ050397 (1-1089)

QY 1 TrrpAlaIleThrValAGluGluArgAlaIshHisAspGlnGlnPheHisSerLeuIlePro 20
DB 292 TGGGCCATTAAGTGAAGAGAGAGAGAGAGAGATATCGAGCTTCCATGTTTAAAGCCA 351
QY 21 IlleSergIlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSergIlyLeuPro 40
DB 352 ATATCTGGATTCATTACGCGTATCAAGCTAGAAAATTTTTCATCTGGGTTACCT 411
QY 41 GlnProValIleuAlaGlnIleTrrpAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
DB 412 CAACCTGTTTAAAGACAGATATGGGCACTAGCATGAAATATGATGAGAAGATGGAT 471
QY 61 GlnValGluPheSerIleAlaMetIleLeuIleLeuIleLeuIleLeuIleGlnIlyTyrGlnLeu 80
DB 472 CAAGTGGAGTTTCCATAGCTATGAAACTTATCAAACTGAAAGCTACAAAGATATCAGCTA 531
QY 81 ProSeraIleuProProValMet 88
DB 532 CCTCTGCACCTTCCCGCTGTCATG 555

RESULT 8
LOCUS BC013578
DEFINITION Homo sapiens, similar to intersectin 1 (SH3 domain protein), clone
IMAGE:3878242, mRNA.
ACCESSION BC013578.1 GI:15488896
VERSION BC013578.1
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2079)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: http://imgc.nci.nih.gov
Contact: MGC help desk
Email: cgabds-r@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.sbcg.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

FEATURES

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAC Plate: 14 Row: C Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796
This clone has the following problem: retained intron.
Location/Qualifiers

ORIGIN

Alignment Scores:
Pred. No.: 1.12e-52 Length: 2079
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BC013578 (1-2079)

QY 1 TrrpAlaIleThrValAGluGluArgAlaIshHisAspGlnGlnPheHisSerLeuIlePro 20
DB 174 TGGGCCATTAAGTGAAGAGAGAGAGAGAGATATCGAGCTTCCATGTTTAAAGCCA 233
QY 21 IlleSergIlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSergIlyLeuPro 40
DB 234 ATATCTGGATTCATTACGCGTATCAAGCTAGAAAATTTTTCATCTGGGTTACCT 293
QY 41 GlnProValIleuAlaGlnIleTrrpAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
DB 294 CAACCTGTTTAAAGACAGATATGGGCACTAGCATGAAATATGATGAGAAGATGGAT 353
QY 61 GlnValGluPheSerIleAlaMetIleLeuIleLeuIleLeuIleLeuIleGlnIlyTyrGlnLeu 80
DB 354 CAAGTGGAGTTTCCATAGCTATGAAACTTATCAAACTGAAAGCTACAAAGATATCAGCTA 413
QY 81 ProSeraIleuProProValMet 88
DB 414 CCTCTGCACCTTCCCGCTGTCATG 437

RESULT 9
LOCUS BC020269
DEFINITION Homo sapiens, clone IMAGE:4899011, mRNA.
ACCESSION BC020269
BC020269 2126 bp mRNA linear HTC 19-DEC-2001

```

VERSION      BC020269.1  GI:17939664
KEYWORDS     HTC.
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 2126)
AUTHORS      Strausberg, R.
TITLE        Direct Submission
JOURNAL      Submitted (19-DEC-2001) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
              Email: gcgabs-r@mail.nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: Rubin Laboratory
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Genome Sequence Centre,
              BC Cancer Agency, Vancouver, BC, Canada
              info@bcsc.bc.ca
              Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
              Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
              Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
              Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
              Ness, Pawan Pandoh, Annd-Lissa Prabhu, Parvaneh Saeedi, Jacqueline
              Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott,
              Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
              George Yang, Scott Zuydam, Marco Marra.

FEATURES
  source      1..2126
              /location=Qualifiers
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:4899011"
              /tissue_type="Pancreas, epithelioid carcinoma"
              /clone_id="NTH MGC 42"
              /lab_host="DH10B-R"
              /note="Vector: pOTB7"

Alignment Scores:
Pred. No.:      1-156-52      Length:      2126
Score:          459.00        Matches:      88
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    100.00%       Indels:        0
DB:             11           Gaps:         0

US-09-720-934-2_COPY_15_102 (1-88) x BC020269 (1-2126)
QY      1 TTPAIAIETHRVALGUGUARGALALYSHISAASPGINGINPHEHISEXLEUYSPRO 20
        |||
Db      224 TGGGCCATTAACGTGAGAGAAAGAGCGAAGCATGATGACGACGTTCCATGTTTAAAGCA 283
QY      21 TLESERGILYPHIELIETHRGILYASPGINLAARGAENPHEPHEGINSERGIYLEUDRO 40
        |||
Db      284 ATATCTGATTCATTAACGTGAGTCAACGCTAGAAAACCTTTTTCATCTGGGTTTACT 343
QY      41 GINPROVALLEUALAGINILIEITPAPALALEUALAASPMETASANSASPGIYARMEVESP 60
        |||
Db      344 CAACCTGTTTTCAGACAGATATGGGCACTAGCTGACATGAATATGATGAGAAAGATGGAT 403
QY      61 GINVALGUPHESERILEALAMECLYSEULEULEYSEULYSEULGINGLYTYRGINDEN 80

```

```

Db      404 CAACTGAGAGTTTCCAGACTATGAAACTTATCAAACTACAGACTACAGATATACGTA 463
QY      81 PROSERIALAUEPROBROVALMET 88
        |||
Db      464 CCTCTGCACTTCCCTCTGTCATG 487

RESULT 10
BU381021
LOCUS
DEFINITION  BU381021 523 bp mRNA linear EST 28-NOV-2002
              60360857F1 CSEQCHN75 Gallus gallus cDNA clone CHEST872a15 5', mRNA
              sequence.
ACCESSION  BU381021.1 GI:25889022
KEYWORDS   EST.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 523)
            Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
            Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
            A Comprehensive Collection of Chicken cDNAs
            Curr. Biol. 12 (22), 1965-1969 (2002)
            2235534
PUBMED     12445392
COMMENT    Contact: Simon Hubbard
            Department of Biomedical Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 0161360409
            Email: Simon.Hubbard@umist.ac.uk.

FEATURES
  source      1..523
              /location=Qualifiers
              /organism="Gallus gallus"
              /mol_type="mRNA"
              /strain="White Leghorn, Hisex"
              /db_xref="taxon:9031"
              /clone="CHEST872a15"
              /dev_stage="36"
              /lab_host="DH10B"
              /clone_id="CSEQCHN75"
              /note="Organ: trunks; Vector: pBluescript II KS(+);
              Site 1: EcoRI; Site 2: NotI; This normalized library was
              constructed from 1 million independent clones. cDNA
              synthesis was initiated using an oligo(dT) primer, using
              methylated C in the first strand synthesis reaction.
              Following this first strand reaction, double-stranded cDNA
              was blunted, ligated to NotI adapters, digested with
              EcoRI, size-selected, and cloned into the NotI and EcoRI
              compatible sites of a custom modified MCS of the
              pBluescript (KS+) vector. The library was normalized in 2
              rounds using conditions adapted from Soares et al., PNAS
              (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
              (1996) 7: 791, except that a significantly longer
              reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.:      2-456-52      Length:      523
Score:          451.00        Matches:      86
Percent Similarity: 98.86%    Conservative: 1
Best Local Similarity: 97.73% Mismatches:      1
Query Match:    98.26%       Indels:        0
DB:             13           Gaps:         0

US-09-720-934-2_COPY_15_102 (1-88) x BU381021 (1-523)
QY      1 TTPAIAIETHRVALGUGUARGALALYSHISAASPGINGINPHEHISEXLEUYSPRO 20
        |||

```

Db 196 TGGGCTATTACTGTGAGGAGAGAGCTAAACATGATCAACAGTTCCACAGTCTGAAACCA 255

QY 21 TlSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40

Db 256 ACATCTGATTTTACATGCTGATCAAGCTAGAAACTTTTTCAGTCTGGGTTACTT 315

QY 41 GlnProValLeuAlaGlnIleThrAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60

Db 316 CAGCGGATGTTAGACAGATATGGGCTTTAGCTGACATGAAACATGATGAGAGATGGAT 375

QY 61 GlnValGlnPheSerIleAlaMetIleValLeuIleIleValLeuGlnGlnGlyTyrGlnLeu 80

Db 376 CAGCTGAGATTTCATAGCTATGAACTTATCAATTAAACTACAGAGCTATCAACTC 435

QY 81 ProSerAlaLeuProProValMet 88

Db 436 CCATCTGCGCTGCTCCTGTCATG 459

RESULT 11

LOCUS BU128229 907 bp mRNA linear EST 25-NOV-2002

DEFINITION 603114019F1 CSEQCHL20 Gallus gallus cDNA clone CHEST6566 5', mRNA

ACCESSION BU128229

VERSION BU128229

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 907)

AUTHORS Boardman,P.E., Sanz-Esguerra,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

TITLE A Comprehensive Collection of Chicken cDNAs

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE 22335534

PUBMED 12445392

COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1..907

Location/Qualifiers

1..907

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="CHEST6566"

/dev_stage="36"

/lab_host="MDH10B"

/clone_lib="CSEQCHL20"

/note="Organ: Limbs; Vector: Bluescript II KS(+); Site: 1: EcoRI; Site 2: NotI; Modification of Bluescript II KS(+) [Stratagene] vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut Bluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BspI and BamHI sites [5'ggccggctgcagcccgagatccgaaataag] [5'aattcttttcgataccggcgacgcg]"

ORIGIN

Alignment Scores:

Pred. No.: 4.99e-52 Length: 907

Score: 451.00 Matches: 86

Percent Similarity: 98.86% Conservative: 1

Best Local Similarity: 97.73% Mismatches: 1

Query Match: 98.26% Indels: 0

DB: 13 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BU128229 (1-907)

QY 1 TlPAlleIleThrValGluGluArgAlaIleHisAspGlnGlnPheHisSerLeuIlePro 20

Db 62 TGGGCTATTACTGTGAGGAGAGAGCTAAACATGATCAACAGTTCCACAGTCTGAAACCA 121

QY 21 TlSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40

Db 122 ACATCTGATTTTACATGCTGATCAAGCTAGAAACTTTTTCAGTCTGGGTTACTT 181

QY 41 GlnProValLeuAlaGlnIleThrAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60

Db 182 CAGCGGATGTTAGACAGATATGGGCTTTAGCTGACATGAAACATGATGAGAGATGGAT 241

QY 61 GlnValGlnPheSerIleAlaMetIleValLeuIleIleValLeuGlnGlnGlyTyrGlnLeu 80

Db 242 CAGCTGAGATTTCATAGCTATGAACTTATCAATTAAACTACAGAGCTATCAACTC 301

QY 81 ProSerAlaLeuProProValMet 88

Db 302 CCATCTGCGCTGCTCCTGTCATG 325

RESULT 12

LOCUS AI852079 445 bp mRNA linear EST 15-JUL-1999

DEFINITION UI-M-BH0-a3a-f-07-0-UI.s1 NIH BMAP M.S1 Mus musculus cDNA clone

ACCESSION AI852079

VERSION AI852079

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 445)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mst@mail.nih.gov
Oligo-dT track not found. Not 1 site shown in beginning of sequence is likely internal to the message. cDNA library Preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH record will be updated accordingly when that means is determined.
Seq primer: M13 Forward
POLYA=No.

FEATURES

source

1..445

Location/Qualifiers

1..445

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH0-a3a-f-07-0-UI"

/dev_stage="27-32 days"

/lab_host="MDH10B (Life Technologies)"

/clone_lib="NIH BMAP M.S1"

/note="Vector: pT7TSB-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH BMAP M.S1 library is a subtracted library derived from a mixture of normalized libraries from ten regions of the

mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 20,000 cDNA clones obtained from non-normalized and normalized libraries of these ten regions of the mouse brain.
TAG: TISSUE=basal-ganglia
TAG: LIB=NIH BMAP_M_S1
TAG_SEQ=GTGAC"

ORIGIN

Alignment Scores:

Pred. No.: 4,93e-51 Length: 445
Score: 441.00 Matches: 85
Percent Similarity: 97.73% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 2
Query Match: 96.08% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x A1852079 (1-445)

QY 1 TTPAIAIEThrValGluGluArgAlaYsHisAspGlnGlnPheHisSerLeuYsPro 20
DB 103 TGGGCCATTAACGTGTGAGAGAAAGGCGCAAGCATGACACGACTTCTTAGCCTGAAGCCG 162
QY 21 ILESerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
DB 163 ATAGCGGATTTATTAAGTGTATCAAGGAGAACTTTTCCATCTGGTTACT 222
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 223 CAGCTGCTTTAGCAACAATATGAGCGCTAGCGCATGAATACGATGAGAGATGAT 282
QY 61 GlnValGluPheSerIleAlaMetYsLeuIleYsLeuYsLeuGlnGlyTyrGlnLeu 80
DB 283 CAAGTGAATTTTCCATGACCATGAGAGCTTATCAACTGAAGCTACAGATATCAGCTC 342
QY 81 ProSerAlaLeuProProValMet 88
DB 343 CCTCCACACTTCCCTGTCATG 366

RESULT 13

LOCUS BE199657 475 bp mRNA linear EST 26-JUN-2000
DEFINITION u970a11.x1 Soares mammary gland NMLMG Mus musculus cDNA clone
IMAGE:1547708 3' similar to TR:Q9ZOR4 Q9ZOR4 ESEIL PROTEIN. ;, mRNA
sequence.

ACCESSION BE199657
VERSION BE199657.1 GI:8711826
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 475)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:953056
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 469.
Location/Qualifiers

FEATURES

source
1..475
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1547708"

/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DB108"
/clone_lib="Soares mammary_gland_NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.: 5.37e-51 Length: 475
Score: 441.00 Matches: 85
Percent Similarity: 97.73% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 2
Query Match: 96.08% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BE199657 (1-475)

QY 1 TTPAIAIEThrValGluGluArgAlaYsHisAspGlnGlnPheHisSerLeuYsPro 20
DB 104 TGGGCCATTAACGTGTGAGAGAAAGGCGCAAGCATGACACGACTTCTTAGCCTGAAGCCG 163
QY 21 ILESerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
DB 164 ATAGCGGATTTATTAAGTGTATCAAGGAGAACTTTTCCATCTGGTTACT 223
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 224 CAGCTGCTTTAGCAACAATATGAGCGCTAGCGCATGAATACGATGAGAGATGAT 283
QY 61 GlnValGluPheSerIleAlaMetYsLeuIleYsLeuYsLeuGlnGlyTyrGlnLeu 80
DB 284 CAAGTGAATTTTCCATGACCATGAGAGCTTATCAACTGAAGCTACAGATATCAGCTC 343
QY 81 ProSerAlaLeuProProValMet 88
DB 344 CCTCCACACTTCCCTGTCATG 367

RESULT 14

LOCUS A1481705 488 bp mRNA linear EST 09-MAR-1999
DEFINITION v930f01.x1 Barstead mouse pooled organs MFLRB4 Mus musculus cDNA
clone IMAGE:888505 3' similar to TR:O42287 O42287 INTERSECTIV. ;,
mRNA sequence.

ACCESSION A1481705
VERSION A1481705.1 GI:4374931
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 488)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Fri Jul 2 10:23:37 2004

Search completed: July 1, 2004, 19:48:26
Job time : 2655.21 secs

us-09-720-934-2_copy_15_102.rst

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 61.6206 Seconds
(without alignments)
792.522 Million cell updates/sec

Title: US-09-720-934-2_COPY_15_102

Perfect score: 459

Sequence: 1 WAIYBERAKHQDQFHSLKP.....KLIKLQGYQLPSALPPVM 88

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delpext 7.0	

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-O/cgnt2_1/USPTO.spool.P/US09720934/runat.30062004.064540.13455/app.query.fasta_1.1386
-DB=Issued Patents NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=numam40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09720934 @CGN 1 1 284 @runat.30062004.064540.13455 -NCPU=6 -ICPU=3
-NO MMAP -LARGEJOURN -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgnt2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgnt2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgnt2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgnt2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgnt2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgnt2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	359	78.2	480 4 US-09-404-879A-60	Sequence 60, Appl
2	359	78.2	480 4 US-09-338-933-60	Sequence 60, Appl
3	359	78.2	480 4 US-09-215-681-60	Sequence 60, Appl
4	359	78.2	480 4 US-09-216-003A-60	Sequence 60, Appl
5	359	78.2	531 4 US-09-404-879A-5	Sequence 5, Appl
6	359	78.2	531 4 US-09-338-933-5	Sequence 5, Appl
7	359	78.2	531 4 US-09-215-681-5	Sequence 5, Appl
8	359	78.2	531 4 US-09-216-003A-5	Sequence 5, Appl
9	359	78.2	2017 4 US-09-404-879A-72	Sequence 72, Appl
10	359	78.2	2017 4 US-09-338-933-72	Sequence 72, Appl
11	359	78.2	2017 4 US-09-215-681-72	Sequence 72, Appl
12	359	78.2	2017 4 US-09-216-003A-72	Sequence 72, Appl

13	205	44.7	3033 1 US-08-095-737-3	Sequence 3, Appl
14	205	44.7	3033 1 US-08-480-145-3	Sequence 3, Appl
15	205	44.7	3033 2 US-08-477-389-3	Sequence 3, Appl
16	203	44.2	4165 1 US-08-095-737-1	Sequence 1, Appl
17	203	44.2	4165 1 US-08-480-145-1	Sequence 1, Appl
18	203	44.2	4165 2 US-08-477-389-1	Sequence 2, Appl
19	199.5	43.5	3348 4 US-09-312-762A-2	Sequence 4, Appl
20	199.5	43.5	14707 4 US-09-312-762A-2	Sequence 4, Appl
21	197.5	43.0	3635 4 US-09-312-762A-6	Sequence 6, Appl
22	190.5	41.5	3508 4 US-09-312-762A-1	Sequence 1, Appl
23	183.5	40.0	2373 4 US-09-023-655-850	Sequence 850, App
24	167.5	36.5	5930 3 US-09-328-111-211	Sequence 211, App
25	163.5	35.6	2815 4 US-09-312-762A-7	Sequence 7, Appl
26	142.5	31.0	5764 4 US-09-312-762A-8	Sequence 8, Appl
27	101	22.0	503 4 US-09-370-838-148	Sequence 148, App
28	98	21.4	845 3 US-08-998-416-543	Sequence 543, App
29	70	15.3	1386 4 US-09-252-991A-3191	Sequence 3191, App
30	70	15.3	1461 4 US-09-252-991A-3000	Sequence 3000, Ap
31	70	15.3	1869 4 US-09-252-991A-2676	Sequence 2676, Ap
32	70	15.3	72928 3 US-09-009-913-1	Sequence 1, Appl
33	69	15.0	592 4 US-09-252-991A-12349	Sequence 12349, A
34	69	15.0	618 4 US-09-252-991A-12199	Sequence 12199, A
35	69	15.0	1026 4 US-09-252-991A-12295	Sequence 12295, A
36	68	14.8	492 3 US-09-328-111-151	Sequence 151, App
37	68	14.8	4091 4 US-09-620-312D-339	Sequence 339, App
38	66.5	14.5	1349 1 US-07-951-715A-20	Sequence 20, Appl
39	66.5	14.5	1349 2 US-08-459-448A-20	Sequence 20, Appl
40	66.5	14.5	1349 3 US-08-459-595A-20	Sequence 20, Appl
41	66.5	14.5	1349 3 US-08-459-504B-20	Sequence 20, Appl
42	66.5	14.5	1349 3 US-08-459-444-20	Sequence 0, Appl
43	66.5	14.5	1349 4 US-09-547-422-20	Sequence 0, Appl
44	63.5	13.8	801 1 US-07-959-946-4	Sequence 4, Appl
45	63.5	13.8	801 1 US-08-333-577-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-404-879A-60
Sequence 60, Application US/09404879A
Patent No. 6468546
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C2
CURRENT APPLICATION NUMBER: US/09/404, 879A
CURRENT FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 393
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 60
LENGTH: 480
TYPE: DNA
ORGANISM: Homo sapien
US-09-404-879A-60

Alignment Scores:

Pred. No.: 1.26e-46 Length: 480
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-404-879A-60 (1-480)

QY 1 Trrpallethrrvalglugluargalalyshisapnglgnphetsserleuyspro 20
DB 56 TGGGCTATTACATCTGAAGAACGTRCTAAGCATGTAACAGTTTGAACTCAAACT 115
QY 21 lleserglyneilrthnglyaspGlnAlaArgasnphetheglnserGlyleuBro 40

```

Db      116 TCAGAGGTTACATACAGGTGATCAAGCCCGTACTTTTCTCTACAGCTCTGCCG 175
QY      41 GlnProValLeuAlaGlnIleTrrPaLeuAlaAspMetAsnAspGlyArgMetAsp 60
Db      176 GCCCGGTTTACCTGGAATATGGGCTTATCAGATCAAGATCAAGATGGGAAGATGGAC 235
QY      61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
Db      236 CAGCAAGAGTTCTCTATACCTATGAACCTCATCAAGTTAAAGTTGCAAGGCCCAACAGCTG 295
QY      81 ProSerAlaLeuProProValMet 88
Db      296 CCTGTAGTCTCCCTCTATCATG 319

RESULT 2
US-09-338-933-60
; Sequence 60, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-60

Alignment Scores:
Pred. No.: 1.26e-46      Length: 480
Score: 359.00           Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21%      Indels: 0
                        Gaps: 0
Db: 4

US-09-720-934-2_COPY_15_102 (1-88) x US-09-338-933-60 (1-480)

QY      1 TrpAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db      56 TGGGCTATTACATCTGGAAGAAAGTACTAAGCATGATTAACACGTTTGATTAACCTCAAACT 115
QY      21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db      116 TCAGAGGTTACATACAGGTGATCAAGCCCGTACTTTTCTCTACAGCTCTGCCG 175
QY      41 GlnProValLeuAlaGlnIleTrrPaLeuAlaAspMetAsnAspGlyArgMetAsp 60
Db      176 GCCCGGTTTACCTGGAATATGGGCTTATCAGATCAAGATCAAGATGGGAAGATGGAC 235
QY      61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
Db      236 CAGCAAGAGTTCTCTATACCTATGAACCTCATCAAGTTAAAGTTGCAAGGCCCAACAGCTG 295
QY      81 ProSerAlaLeuProProValMet 88
Db      296 CCTGTAGTCTCCCTCTATCATG 319

RESULT 3
US-09-215-681-60
; Sequence 60, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.

```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-60

Alignment Scores:
Pred. No.: 1.26e-46      Length: 480
Score: 359.00           Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21%      Indels: 0
                        Gaps: 0
Db: 4

US-09-720-934-2_COPY_15_102 (1-88) x US-09-215-681-60 (1-480)

QY      1 TrpAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db      56 TGGGCTATTACATCTGGAAGAAAGTACTAAGCATGATTAACACGTTTGATTAACCTCAAACT 115
QY      21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db      116 TCAGAGGTTACATACAGGTGATCAAGCCCGTACTTTTCTCTACAGCTCTGCCG 175
QY      41 GlnProValLeuAlaGlnIleTrrPaLeuAlaAspMetAsnAspGlyArgMetAsp 60
Db      176 GCCCGGTTTACCTGGAATATGGGCTTATCAGATCAAGATCAAGATGGGAAGATGGAC 235
QY      61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
Db      236 CAGCAAGAGTTCTCTATACCTATGAACCTCATCAAGTTAAAGTTGCAAGGCCCAACAGCTG 295
QY      81 ProSerAlaLeuProProValMet 88
Db      296 CCTGTAGTCTCCCTCTATCATG 319

RESULT 4
US-09-216-003A-60
; Sequence 60, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-60

Alignment Scores:
Pred. No.: 1.26e-46      Length: 480
Score: 359.00           Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21%      Indels: 0
                        Gaps: 0
Db: 4

US-09-720-934-2_COPY_15_102 (1-88) x US-09-216-003A-60 (1-480)

```

```
QY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
    |||||
DB 56 TGGGCTATTACATGTAAGAAAGCTACTAAGCATGATGAACAGTTTGATTAACCTCAAACT 115
    |||||
QY 21 ILeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
    |||||
DB 116 TCAGAGAGTTACATTAACAGGTGATCAAGCCCGTACTCTTTTCCACAGCTCAGGCTCCG 175
    |||||
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAspGlyTyrGlnLeu 60
    |||||
DB 176 GCCCGGTTTAGCTGAATAATGGCCCTTACAGATCTGAACAGAGATGGGAGATGAC 235
    |||||
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysGlnGlyTyrGlnLeu 80
    |||||
DB 236 CAGCAAGAGTTCTTATAGCTATGAACCTCATCAAGTTAAAGTTGACAGGGCCAAAGCTG 295
    |||||
QY 81 ProSerAlaLeuProProValMet 88
    |||||
DB 296 CCTGTAGTCTCTCCCTCATCATG 319
    |||||

RESULT 5
US-09-404-879A-5
; Sequence 5, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-5

Alignment Scores:
Pred. No.: 1,46e-46 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-404-879A-5 (1-531)
QY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
    |||||
DB 84 TGGGCTATTACATGTAAGAAAGCTACTAAGCATGATGAACAGTTTGATTAACCTCAAACT 143
    |||||
QY 21 ILeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
    |||||
DB 144 TCAGAGAGTTACATTAACAGGTGATCAAGCCCGTACTCTTTTCCACAGCTCAGGCTCCG 203
    |||||
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAspGlyTyrGlnLeu 60
    |||||
DB 204 GCCCGGTTTAGCTGAATAATGGCCCTTATCAAGTCTGAACAGAGATGGGAGATGAC 263
    |||||
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysGlnGlyTyrGlnLeu 80
    |||||
DB 264 CAGCAAGAGTTCTTATAGCTATGAACCTCATCAAGTTAAAGTTGACAGGGCCAAAGCTG 323
    |||||
QY 81 ProSerAlaLeuProProValMet 88
    |||||
DB 324 CCTGTAGTCTCTCCCTCATCATG 347
    |||||

RESULT 6
US-09-338-933-5
```

```
; Sequence 5, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-5

Alignment Scores:
Pred. No.: 1,46e-46 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-338-933-5 (1-531)
QY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
    |||||
DB 84 TGGGCTATTACATGTAAGAAAGCTACTAAGCATGATGAACAGTTTGATTAACCTCAAACT 143
    |||||
QY 21 ILeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
    |||||
DB 144 TCAGAGAGTTACATTAACAGGTGATCAAGCCCGTACTCTTTTCCACAGCTCAGGCTCCG 203
    |||||
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAspGlyTyrGlnLeu 60
    |||||
DB 204 GCCCGGTTTAGCTGAATAATGGCCCTTATCAAGTCTGAACAGAGATGGGAGATGAC 263
    |||||
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysGlnGlyTyrGlnLeu 80
    |||||
DB 264 CAGCAAGAGTTCTTATAGCTATGAACCTCATCAAGTTAAAGTTGACAGGGCCAAAGCTG 323
    |||||
QY 81 ProSerAlaLeuProProValMet 88
    |||||
DB 324 CCTGTAGTCTCTCCCTCATCATG 347
    |||||

RESULT 7
US-09-215-681-5
; Sequence 5, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-5

Alignment Scores:
Pred. No.: 1,46e-46 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
```


Alignment Scores:	
Pred. No.:	1.02e-45
Score:	359.00
Percent Similarity:	85.23%
Best Local Similarity:	76.14%
Query Match:	78.21%
DB:	4
Length:	2017
Matches:	67
Conservative:	8
Mismatches:	13
Indels:	0
Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-338-933-72 (1-2017)

[illegible]

RESULT 11
US-09-215-681-72
; Sequence 72, Application US/09215681A

```

; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF OVARIAN CANCER
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-72

```

[illegible]

US-09-720-934-2_COPY_15_102 (1-88) X US-09-215-681-72 (1-2017)

QY 1 TPRLAILEthrya]GluGluArgAlaIysHisAspGlnGlnPheHisSerLeuIysPro 20

Db 84 TGGCGATTATACACTGTGAAGACGTACTAACATGATATAACAGTTTGATACCTTCAACT 143

QY 21 ILeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40

Db 144 TCGAGAGGTTACATCAACAGGTGATCAAGCCCGTACTTTTCTACAGTCAAGGCTCCCG 203

QY 41 GlnProValIleuAlaGlnIleTPRLAILEuAlaAspMetAsnAsnAspGlyArgMetAsp 60

Db 204 GCCCCGGTTTTCGTGAATATATGGCGCTTTATCAAGATTCGAACAAAGATGGGAAATGGAC 263

Qy	61	GIlnaIGluPheSerIleIaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu	80
Db	264	CAGCAAGATTTCTTATAGCTATGAAACATCATCAATTAAGTTCAGAGGCCAACAGCTG	323
Qy	81	ProSerLeuIleuProProValMet	88
Db	324	CGTATAGTCTCTCCCTCTATCATGATG	347

RESULT 12
ME-09-316-003A-73

```

; Sequence 72, Application US/09216003A
; Patent No. 6670463
;
; GENERAL INFORMATION:
;
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 72
;
; LENGTH: 2017
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
US-09-216-003A-72

```

Alignment Scores:

Pred. No.:	1.02e-45	Length:	2017
Score:	359.00	Matches:	67
Percent Similarity:	85.23%	Conservative:	8
Best Local Similarity:	76.14%	Mismatches:	13
Query Match:	78.21%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) X US-09-216-003A-72 (1-2017)

QY 1TPPa1a1eThVa1g1u1g1u1r1g1a1a1y1S1H1s1a1S1p1G1n1p1h1e1S1e1r1L1e1u1S1P1r1o 20
Db 1TGGGCTATTACATCTGAAAGAACCTACTAAGACATGATAAACAGTTTGATATACCTCAAACT 144
QY 211e1S1e1r1G1y1P1h1e11e1Th1r1G1y1A1S1p1G1n1a1a1a1r1G1A1S1n1P1h1e1D1h1e1G1n1S1e1r1G1y1L1e1u1P1r1o 40
Db 144TCAGGAGGTTACATACAGATGATCAAGCCCGTACTTTTTCCTAAGTAGGTCTGCGC 203
QY 41G1n1P1r1o1V1a1l1e1u1a1a1n11e1T1r1P1a1a1e1u1a1a1S1P1M1e1r1A1S1n1S1n1S1P1G1Y1A1G1M1e1r1A1S1P 60
Db 204GCCCGGTTTGTGCTAAAAATGCGGCTTATCAATCTGAACAAAGATGGGAAAGATGGAC 265
QY 61G1n1V1a1l1G1u1P1h1e1S1e1r11e1a1M1e1T1L1e1u11e1L1e1u1S1e1u1S1e1u1G1n1G1Y1r1G1n1L1e1u 80
Db 264CAGCAAGAGTTCTCTTAGCTATGAACATCATCAAGTTAAAGTTGACAGGCGCAACAGCTG 322
QY 81P1r1o1S1e1r1A1L1e1u1P1r1o1P1r1o1V1a1l1M1e1r 88
Db 324CCTGAGTCTCTCTCTATCAATG 347

RESULT 13
TIC 06 005-737-3

```

: Sequence 3, Application US/08095737
: Patent No. 5487979
:
: GENERAL INFORMATION:
:
: APPLICANT: DIFIORE, Pier P
:
: APPLICANT: Fazio, Francesca
:
: TITLE OF INVENTION: A Substrate for the Epidermal Growth
:
: FACTOR OF INVENTION: Factor Receptor Kinase
:
: NUMBER OF SEQUENCES: 4
:
: CORRESPONDENCE ADDRESSES:
:
: ADDRESSEE: Knobbe, Martens, Olson & Bear
:
: STREET: 620 Newport Center Drive, Sixteenth Floor
:
: CITY: Newport Beach
:

```

STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,737
FILING DATE: 1993/07/22
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-095-737-3

Alignment Scores:
Pred. No.: 1,75e-21 Length: 3033
Score: 205.00 Matches: 34
Percent Similarity: 68.18% Conservative: 26
Best Local Similarity: 38.64% Mismatches: 28
Query Match: 44.66% Indels: 0
DB: 1 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-08-095-737-3 (1-3033)

QY 1 TTPAAlIeThVaGlUGlUaYrGAlaYsHsAspGInGlnPheHIsSerLeuYsPro 20
Db 474 TGGGCTGTAAGCTCGAAGATTAAGCAATATGATGCAATTTTGCACGTTTAAGCCCA 533
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArGAsnPhePheGlnSerGlyLeuPro 40
Db 534 GTGATGATGATTTTGTCTGTGATTAAGCAACAGTGTGCTCAACTCAAGTTTACT 593
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaaspMetAsnAspGlyArGmetAap 60
Db 594 GTGGAATCTCTTGAAGATTGGAGATTGAGTGAATATGACACAGATGAAAGCTGGAC 653
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 654 AGAGATGAGTTTGACGTTGCCATGTTTGTATACTGTGACGTGAGAAAGAACCTGTG 713
QY 81 ProSerAlaLeuProProValMet 88
Db 714 CCAATGCTCTTCCCTCCAGCCTTG 737

RESULT 14
US-08-480-145-3
Sequence 3, Application US/08480145
Patent No. 5717067
GENERAL INFORMATION:
APPLICANT: DiFiore, Pier P
APPLICANT: Fazio, Francesca
TITLE OF INVENTION: A Substrate for the Epidermal Growth
Factor Receptor Kinase
NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,145
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/095,737
FILING DATE: 22-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-480-145-3

Alignment Scores:
Pred. No.: 1,75e-21 Length: 3033
Score: 205.00 Matches: 34
Percent Similarity: 68.18% Conservative: 26
Best Local Similarity: 38.64% Mismatches: 28
Query Match: 44.66% Indels: 0
DB: 1 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-08-480-145-3 (1-3033)

QY 1 TTPAAlIeThVaGlUGlUaYrGAlaYsHsAspGInGlnPheHIsSerLeuYsPro 20
Db 474 TGGGCTGTAAGCTCGAAGATTAAGCAATATGATGCAATTTTGCACGTTTAAGCCCA 533
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArGAsnPhePheGlnSerGlyLeuPro 40
Db 534 GTGATGATGATTTTGTCTGTGATTAAGCAACAGTGTGCTCAACTCAAGTTTACT 593
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaaspMetAsnAspGlyArGmetAap 60
Db 594 GTGGAATCTCTTGAAGATTGGAGATTGAGTGAATATGACACAGATGAAAGCTGGAC 653
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 654 AGAGATGAGTTTGACGTTGCCATGTTTGTATACTGTGACGTGAGAAAGAACCTGTG 713
QY 81 ProSerAlaLeuProProValMet 88
Db 714 CCAATGCTCTTCCCTCCAGCCTTG 737

RESULT 15
US-08-477-389-3
Sequence 3, Application US/08477389

1 Patent No. 5872219
2
3 GENERAL INFORMATION:
4 APPLICANT: DiFiore, Pier P
5 APPLICANT: Fazio, Francesca
6 TITLE OF INVENTION: A Substrate for the Epidermal Growth
7 TITLE OF INVENTION: Factor Receptor Kinase
8 NUMBER OF SEQUENCES: 4
9
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Knobb, Martens, Olson & Bear
12 STREET: 620 Newport Center Drive, Sixteenth Floor
13 CITY: Newport Beach
14 STATE: California
15 COUNTRY: United States of America
16 ZIP: 92660
17
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: PatentIn Release #1.0, Version #1.25
23
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/08/477,389
26 FILING DATE: 07-JUN-1995
27
28 CLASSIFICATION: 530
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: 08/095,737
31 FILING DATE: 22-JUN-1993
32 ATTORNEY/AGENT INFORMATION:
33 NAME: Israelsen, Ned A
34 REGISTRATION NUMBER: 29,655
35 REFERENCE/DOCKET NUMBER: NH060.001A
36
37 TELECOMMUNICATION INFORMATION:
38 TELEPHONE: (619) 235-8550
39 TELEFAX: (619) 235-0176
40
41 INFORMATION FOR SEQ ID NO: 3:
42 SEQUENCE CHARACTERISTICS:
43 LENGTH: 303 base pairs
44 TYPE: nucleic acid
45 STRANDEDNESS: double
46 TOPOLOGY: linear
47
48 MOLECULE TYPE: cDNA
49 HYPOTHETICAL: NO
50 ANTI-SENSE: NO
51
52 FEATURE:
53 NAME/KEY: CDS
54 LOCATION: 111..2802
55
56 US-08-477-389-3

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 269.977 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_15_102
Perfect score: 459
Sequence: 1 WAITBERAHDQGFSLKP.....KLTKLQGYQLPSALPPVM 88

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Rgapop 6.0 , Rgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3163042 segs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/p/US09720934/runat.30062004.064541.13481/app.query.fasta_1.1386
-DB=Published Applications NA -OPM=fastat -SUFFIX=nmdb -MINMATCH=0.1
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62
-TRANS=numat40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09720934 @CGN 1 1 1500 @runat.30062004.064541_13481
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAOP=6 -FGAEXT=7 -YGAOP=10 -YGAEXT=0.5 -DELDP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10C_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

1	459	100.0	3319	11	US-09-764-875-88	Sequence 88, Appl
2	459	100.0	3466	16	US-10-158-057-33	Sequence 33, Appl
3	425.5	92.7	568	10	US-09-764-881-55	Sequence 55, Appl
4	425.5	92.7	568	11	US-09-764-875-404	Sequence 404, Appl
5	425.5	92.7	568	13	US-09-764-881-55	Sequence 55, Appl
6	425.5	92.7	568	16	US-10-242-747-55	Sequence 55, Appl
7	425.5	92.7	568	16	US-10-158-057-127	Sequence 127, Appl
8	365	79.5	5828	13	US-10-398-8854-15	Sequence 12, Appl
9	359	78.2	480	9	US-09-884-441-60	Sequence 60, Appl
10	359	78.2	480	10	US-09-907-969-60	Sequence 60, Appl
11	359	78.2	480	10	US-09-827-271-60	Sequence 60, Appl
12	359	78.2	531	15	US-10-198-053-60	Sequence 60, Appl
13	359	78.2	531	9	US-09-884-441-55	Sequence 5, Appl
14	359	78.2	531	10	US-09-907-969-5	Sequence 5, Appl
15	359	78.2	531	10	US-09-827-271-5	Sequence 5, Appl
16	359	78.2	531	15	US-10-198-053-5	Sequence 5, Appl
17	359	78.2	2017	9	US-09-884-441-72	Sequence 72, Appl
18	359	78.2	2017	10	US-09-907-969-72	Sequence 72, Appl
19	359	78.2	2017	10	US-09-827-271-72	Sequence 72, Appl
20	359	78.2	2017	15	US-10-198-053-72	Sequence 72, Appl
21	351	76.5	4662	10	US-09-198-053-72	Sequence 21728, A
22	339	73.9	967	10	US-09-764-881-50	Sequence 50, Appl
23	339	73.9	967	11	US-09-764-875-411	Sequence 411, Appl
24	339	73.9	967	13	US-09-764-881-50	Sequence 50, Appl
25	339	73.9	967	16	US-10-242-747-50	Sequence 50, Appl
26	339	73.9	967	16	US-10-158-057-40	Sequence 40, Appl
27	212	46.2	292	9	US-09-864-761-20261	Sequence 20261, A
28	212	46.2	304	9	US-09-864-761-21373	Sequence 21373, A
29	212	46.2	310	9	US-09-864-761-19751	Sequence 19751, A
30	212	46.2	310	9	US-09-864-761-19759	Sequence 19759, A
31	212	46.2	466	9	US-09-864-761-3491	Sequence 3491, Ap
32	212	46.2	466	9	US-09-864-761-4631	Sequence 4631, Ap
33	199.5	43.5	605	12	US-10-152-119A-1126	Sequence 1126, Ap
34	199.5	43.5	860	13	US-10-027-632-165796	Sequence 165796, A
35	199.5	43.5	860	13	US-10-027-632-165797	Sequence 165797, A
36	199.5	43.5	860	16	US-10-027-632-165796	Sequence 165796, A
37	199.5	43.5	860	16	US-10-027-632-165797	Sequence 165797, A
38	199.5	43.5	3348	9	US-09-312-762A-2	Sequence 2, Appl
39	199.5	43.5	3485	15	US-10-171-581-354	Sequence 354, App
40	199.5	43.5	14707	9	US-09-312-762A-3	Sequence 3, Appl
41	198	43.1	662	13	US-10-027-632-290824	Sequence 290824, A
42	198	43.1	662	16	US-10-027-632-290824	Sequence 290824, A
43	197.5	43.0	3635	9	US-09-312-762A-6	Sequence 6, Appl
44	193.5	42.2	551	15	US-10-029-386-25205	Sequence 25205, A
45	190.5	41.5	3508	9	US-09-312-762A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-764-875-88
; Sequence 88, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 3319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-875-88

Alignment Scores: 1.41e-62 Length: 3319
Pred. No.: 459.00 Matches: 88
Score: 100.00% Conservative: 0
Percent Similarity: 100.00%

```
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 11
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-09-764-875-88 (1-3319)

QY 1 TTPAAlleThrValGluGlnArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 201 TGGGCGATTAAGTCTAGAGAAAGAGGAGCATGATCAGAGTTCCATGTTAAAGCCA 260

QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 261 ATATCTGATTCATTAATCTGCTGATCAAGCTAGAACTTTTTCATCTGGATTACCT 320

QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 321 CAACCTGTTTACACAGATATGGGCACTGACATGATATATATGGAAGATGAT 380

QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrglnLeu 80
DB 381 CAAGTGAAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGGATATCAGCTA 440

QY 81 ProSerAlaLeuProProValMet 88
DB 441 CCTCTGCACTTCCCTCTGTCAAG 464

RESULT 2
US-10-158-057-33
; Sequence 33, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ205C1
; CURRENT APPLICATION NUMBER: US/10/158,057
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 3466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3194)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3465)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-158-057-33

Alignment Scores:
Pred. No.: 1.5e-62 Length: 3466
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-10-158-057-33 (1-3466)

QY 1 TTPAAlleThrValGluGlnArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 319 TGGGCGATTAAGTCTAGAGAAAGAGGAGCATGATCAGAGTTCCATGTTAAAGCCA 378

QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 379 ATATCTGATTCATTAATCTGCTGATCAAGCTAGAACTTTTTCATCTGGATTACCT 438

QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 441 CCTCTGCACTTCCCTCTGTCAAG 464
```

```
DB 439 CAACCTGTTTACACAGATATGGGCACTGACATGATATATGATGGAAGATGAT 498
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrglnLeu 80
DB 499 CAAGTGAAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGGATATCAGCTA 558

QY 81 ProSerAlaLeuProProValMet 88
DB 559 CCTCTGCACTTCCCTCTGTCAAG 582

RESULT 3
US-09-764-881-55
; Sequence 55, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; PRIOR APPLICATION DATA REMOVED - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-55

Alignment Scores:
Pred. No.: 2.81e-58 Length: 568
Score: 425.50 Matches: 87
Percent Similarity: 98.86% Conservative: 0
Best Local Similarity: 98.86% Mismatches: 1
Query Match: 92.70% Indels: 1
DB: 10 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-09-764-881-55 (1-568)

QY 1 TTPAAlleThrValGluGlnArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 120 TGGGCGATTAAGTCTAGAGAAAGAGGCA--CATGATCAGACATTCATGTTAAAGCCA 177

QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 178 ATATCTGATTCATTAATCTGCTGATCAAGCTAGAACTTTTTCATCTGGATTACCT 237

QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 238 CAACCTGTTTACACAGATATGGGCACTGACATGATATATGATGGAAGATGAT 297

QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrglnLeu 80
DB 298 CAAGTGAAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGGATATCAGCTA 357

QY 81 ProSerAlaLeuProProValMet 88
DB 358 CCTCTGCACTTCCCTCTGTCAAG 381

RESULT 4
```

```
US-09-764-875-404
; Sequence 404, Application US/09764875
; Publication No. US2004001869A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 404
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-404

Alignment Scores:
Pred. No.: 2,81e-58 Length: 568
Score: 425.50 Matches: 87
Percent Similarity: 98.86% Conservative: 0
Best Local Similarity: 98.86% Mismatches: 1
Query Match: 92.70% Indels: 1
Db: 11 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-764-875-404 (1-568)
QY 1 TrrpAlaIeThrrVAlGluGluArgAlaIySHsAspGInGInPheHISerLeuIySPro 20
Db 120 TGGGCGATTAAGTGTAGAGAGAAAGACGA--CATGATCGAGCTTCCATAGTTTAAAGCA 177
QY 21 IleserGIyPheIleThrGIyAspGInAlaArgAsnPhenPheGInSerGIyLeuPro 40
Db 178 ATATCTGATTCATTACGTGGTATCAAGCTAGAACTTTTTCATCTGGGTTACT 237
QY 41 GlnProValIleuAlaGInIleTrrpAlaIeAlaAspMetAsnAspGIyArgMetAsp 60
Db 238 CAACCTGTTTATAGACAGATATGGGCACTAGCTGACATGAATATATGATGGAAGATGAT 297
QY 61 GlnValGluPheSerIleAlaMetIySLeuIleIySLeuIySLeuGInGIyTySInleu 80
Db 298 CAAGTGGAGTTTCCATAGCTATGAACTTATCAACTGAACTGCAAGGATATACGCTA 357
QY 81 ProSerAlaIeuProIvalMet 88
Db 358 CCTCTGCACCTTCCCCCTGTGATG 381

RESULT 5
US-09-764-881-55
; Sequence 55, Application US/09764881
; Publication No. US20020086821A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
```

```
SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-55

Alignment Scores:
Pred. No.: 2,81e-58 Length: 568
Score: 425.50 Matches: 87
Percent Similarity: 98.86% Conservative: 0
Best Local Similarity: 98.86% Mismatches: 1
Query Match: 92.70% Indels: 1
Db: 13 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-764-881-55 (1-568)
QY 1 TrrpAlaIeThrrVAlGluGluArgAlaIySHsAspGInGInPheHISerLeuIySPro 20
Db 120 TGGGCGATTAAGTGTAGAGAGAAAGACGA--CATGATCGAGCTTCCATAGTTTAAAGCA 177
QY 21 IleserGIyPheIleThrGIyAspGInAlaArgAsnPhenPheGInSerGIyLeuPro 40
Db 178 ATATCTGATTCATTACGTGGTATCAAGCTAGAACTTTTTCATCTGGGTTACT 237
QY 41 GlnProValIleuAlaGInIleTrrpAlaIeAlaAspMetAsnAspGIyArgMetAsp 60
Db 238 CAACCTGTTTATAGACAGATATGGGCACTAGCTGACATGAATATATGATGGAAGATGAT 297
QY 61 GlnValGluPheSerIleAlaMetIySLeuIleIySLeuIySLeuGInGIyTySInleu 80
Db 298 CAAGTGGAGTTTCCATAGCTATGAACTTATCAACTGAACTGCAAGGATATACGCTA 357
QY 81 ProSerAlaIeuProIvalMet 88
Db 358 CCTCTGCACCTTCCCCCTGTGATG 381

RESULT 6
US-10-242-747-55
; Sequence 55, Application US/10242747
; Publication No. US20040005577A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ207C1
; CURRENT APPLICATION NUMBER: US/10/242,747
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,881
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
```

```

; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 55
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-242-747-55

```

Alignment Scores:

```

Pred. No.: 2,81e-58
Score: 425.50
Percent Similarity: 98.868
Best Local Similarity: 98.868
Query Match: 92.70%
DB: 16
Length: 568
Matches: 87
Conservative: 0
Mismatch: 1
Indels: 1
Gaps: 0

```

US-09-720-934-2_COPY_15_102 (1-88) * US-10-242-747-55 (1-568)

```

QY 1 TTPAIAIEthVAlGluGluArgAlaYSHsAspGlnGlnPheHisSerLeuYsPro 20
DB 120 TGGGCGCATTAAGTGTAGAGAAAGCGCA--CATGATCAGCAGTTCCATGTTAAAGCCA 177
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 178 ATATCTGATTAATCACTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTACCT 237
QY 41 GlnProValLeuAlaGlnIleTTPAIAIEthVAlGluGluArgAlaYSHsAspGlnGlnPheHisSerLeuYsPro 60
DB 238 CAACCTGTTTTCAGCAGATATGGGCACTAGCTGACATATATGATGAAAGATGAT 297
QY 61 GlnValGlnPheSerIleAlaMetIleYsLeuYsLeuGlnGlyTyrglnLeu 80
DB 298 CAAGTGGAGTTTTCATTAAGAACTTATCAAACTGAACTGAAAGATATACGCTA 357
QY 81 ProSerAlaLeuProProValMet 88
DB 358 CCTCTGCACTTCCCTGCTCATG 381

```

RESULT 7

```

US-10-158-057-127
; Sequence 127, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1205C1
; CURRENT APPLICATION NUMBER: US/10/158, 057
; PRIOR FILING DATE: 2002-06-12
; Prior Application removed - See File Wrapper or Palm

```

```

; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 127
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-158-057-127

```

Alignment Scores:

```

Pred. No.: 2,81e-58
Score: 425.50
Percent Similarity: 98.868
Best Local Similarity: 98.868
Query Match: 92.70%
DB: 16
Length: 568
Matches: 87
Conservative: 0
Mismatch: 1
Indels: 1
Gaps: 0

```

US-09-720-934-2_COPY_15_102 (1-88) * US-10-158-057-127 (1-568)

```

QY 1 TTPAIAIEthVAlGluGluArgAlaYSHsAspGlnGlnPheHisSerLeuYsPro 20
DB 120 TGGGCGCATTAAGTGTAGAGAAAGCGCA--CATGATCAGCAGTTCCATGTTAAAGCCA 177
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 178 ATATCTGATTAATCACTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTACCT 237
QY 41 GlnProValLeuAlaGlnIleTTPAIAIEthVAlGluGluArgAlaYSHsAspGlnGlnPheHisSerLeuYsPro 60
DB 238 CAACCTGTTTTCAGCAGATATGGGCACTAGCTGACATATATGATGAAAGATGAT 297
QY 61 GlnValGlnPheSerIleAlaMetIleYsLeuYsLeuGlnGlyTyrglnLeu 80
DB 298 CAAGTGGAGTTTTCATTAAGAACTTATCAAACTGAACTGAAAGATATACGCTA 357
QY 81 ProSerAlaLeuProProValMet 88
DB 358 CCTCTGCACTTCCCTGCTCATG 381

```

RESULT 8

```

US-10-398-885A-15
; Sequence 15, Application US/10398885A
; Publication No. US20040053282A1
; GENERAL INFORMATION:
; APPLICANT: Sugita, Yuji
; APPLICANT: Hashida, Ryotchi
; APPLICANT: Ogawa, Kaeoru
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Obayashi, Masaya
; APPLICANT: Saito, Hirohisa
; APPLICANT: Takahashi, Eiki
; TITLE OF INVENTION: Method of Testing For Allergic Diseases
; FILE REFERENCE: SHIMIZU-07907
; CURRENT APPLICATION NUMBER: US/10/398, 885A
; PRIOR FILING DATE: 2003-08-11
; CURRENT APPLICATION NUMBER: PCT/JP01/08937
; PRIOR FILING DATE: 2001-10-11
; Prior Application removed - See File Wrapper or Palm

```

PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 5828
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
NAME/KEY: CDS
LOCATION: (7)..(5052)
OTHER INFORMATION:
US-10-398-865A-15

Alignment Scores:
Pred. No.: 3,89e-47 Length: 5828
Score: 365.00 Matches: 68
Percent Similarity: 86.36% Conservative: 8
Best Local Similarity: 77.27% Mismatches: 12
Query Match: 79.52% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-10-398-865A-15 (1-5828)

QY 1 TTPAlalleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 85 TGGGCTATTACCTCTGAAGAACGTAAGCATGACAGCGTTGATTAACCTAAACCT 144
QY 21 TLeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 145 TCAGAGAGTTACATACACAGGATGACAGCGTAACTTTTCCATCAATCAGGCTGCGC 204
QY 41 GlnProValLeuAlaGlnIleTyrAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
DB 205 GCCCTGTTTAGCTGAATATATGGGCTTATCAGACCTAAACAAGATGGGAAGATGAT 264
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 265 CAGCAAGAGTTCTCATACATGATGATAAAGCTCAAGGCTTCAAGGCCAACAGTTG 324
QY 81 ProSerAlaLeuProProValMet 88
DB 325 CCTGTGTTCTCCCTCATATATG 348

RESULT 9
US-09-884-441-60
Sequence 60, Application US/09884441
Patent No. US20020119158A1
GENERAL INFORMATION:
APPLICANT: Algate, Paul A.
APPLICANT: Carter, Patrick
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C7
CURRENT APPLICATION NUMBER: US/09/884,441
CURRENT FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 489
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 60
LENGTH: 480
TYPE: DNA
ORGANISM: Homo sapien
US-09-884-441-60

Alignment Scores:
Pred. No.: 1.06e-47 Length: 480
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-884-441-60 (1-480)

QY 1 TTPAlalleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 56 TGGGCTATTACATCTGAAGAACGTAAGCATGATAAAGCTTTGATTAACCTAAACCT 115
QY 21 TLeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 116 TCAGAGAGTTACATACACAGGATGACAGCGTAACTTTTCCATCAATCAGGCTGCGC 175
QY 41 GlnProValLeuAlaGlnIleTyrAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
DB 176 GCCCGGTTTAGCTGAATATATGGGCTTATCAGACCTTCAAGCAAGATGGGAAGATGAC 235
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 236 CAGCAAGAGTTCTCATACATGATGATAAAGCTCAAGTTAAAGTTCAGAGGCCAACAGCTG 295
QY 81 ProSerAlaLeuProProValMet 88
DB 296 CCTGTGTTCTCCCTCATATATG 319

RESULT 10
US-09-907-969-60
Sequence 60, Application US/09907969
Publication No. US20030091580A1
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
APPLICANT: Retter, Marc W.
APPLICANT: Fanger, Gary Richard
APPLICANT: Reed, Steven G.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Patrick
APPLICANT: Hill, Paul
APPLICANT: Albone, Barl
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.462C8
CURRENT APPLICATION NUMBER: US/09/907,969
CURRENT FILING DATE: 2001-07-17
NUMBER OF SEQ ID NOS: 596
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 60
LENGTH: 480
TYPE: DNA
ORGANISM: Homo sapiens
US-09-907-969-60

Alignment Scores:
Pred. No.: 1.06e-47 Length: 480
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-907-969-60 (1-480)

QY 1 TTPAlalleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 56 TGGGCTATTACATCTGAAGAACGTAAGCATGATAAAGCTTTGATTAACCTAAACCT 115
QY 21 TLeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 116 TCAGAGAGTTACATACACAGGATGACAGCGTAACTTTTCCATCAATCAGGCTGCGC 175
QY 41 GlnProValLeuAlaGlnIleTyrAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
DB 176 GCCCGGTTTAGCTGAATATATGGGCTTATCAGACCTTCAAGCAAGATGGGAAGATGAC 235

```

OY 61 GlnValIleThrValAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
DB 236 CAGCAAGAGTTCTCTAATGCTATGAACCTCATCAAGTTAAAGTTGAGGCCAACAGCTG 295
OY 81 ProSerAlaLeuProProValMet 88
DB 296 CCTGTAGTCCTCCCTCATCATG 319

RESULT 11
US-09-827-271-60
; Sequence 60, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C6
; CURRENT APPLICATION NUMBER: US/09/827,271
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-827-271-60

Alignment Scores:
Pred. No.: 1.06e-47 Length: 480
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-827-271-60 (1-480)

OY 1 TPpAlaIleThrValGlnGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 56 TGGGCTATTACATCTGAAGAAAGTCTAAGCATGATAAACAAGTTGATTAAGCTCAAACT 115
OY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
DB 116 TCAGAGAGTTATACATACAGGCTATCAAGCCCTACTTTTCTTACAGTCAGTCTGCCG 175
OY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 176 GCCCGGTTTATGCTGAATAATAGGCGCTTATCAGATCTGAACAAGAGATGGAAAGATGGAC 235
OY 61 GlnValIleThrValAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
DB 236 CAGCAAGAGTTCTCTAATGCTATGAACCTCATCAAGTTAAAGTTGAGGCCAACAGCTG 295
OY 81 ProSerAlaLeuProProValMet 88
DB 296 CCTGTAGTCCTCCCTCATCATG 319

RESULT 12
US-10-198-053-60
; Sequence 60, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624

```

```

; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-053-60

Alignment Scores:
Pred. No.: 1.06e-47 Length: 480
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-10-198-053-60 (1-480)

OY 1 TPpAlaIleThrValGlnGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 56 TGGGCTATTACATCTGAAGAAAGTCTAAGCATGATAAACAAGTTGATTAAGCTCAAACT 115
OY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
DB 116 TCAGAGAGTTATACATACAGGCTATCAAGCCCTACTTTTCTTACAGTCAGTCTGCCG 175
OY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 176 GCCCGGTTTATGCTGAATAATAGGCGCTTATCAGATCTGAACAAGAGATGGAAAGATGGAC 235
OY 61 GlnValIleThrValAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
DB 236 CAGCAAGAGTTCTCTAATGCTATGAACCTCATCAAGTTAAAGTTGAGGCCAACAGCTG 295
OY 81 ProSerAlaLeuProProValMet 88
DB 296 CCTGTAGTCCTCCCTCATCATG 319

RESULT 13
US-09-884-441-5
; Sequence 5, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Derrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-5

Alignment Scores:
Pred. No.: 1.22e-47 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-884-441-5 (1-531)

OY 1 TPpAlaIleThrValGlnGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 84 TGGGCTATTACATCTGAAGAAAGTCTAAGCATGATAAACAAGTTGATTAAGCTCAAACT 143
OY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40

```

```
Db 144 TCAGAGGTTACATTAACAGGTGATCAAGCCCGTACTTTTTTCTACAGTCAAGGTCGCG 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetLysnAspGlyArgMetAsp 60
Db 204 GCCCGGTTTACTTAAATATGGCCCTTATCAGATCTGAACAGATGGAGATGGAC 263
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysGlnGlyTyrGlnLeu 80
Db 264 CAGCAAGAGTTCTCTATAGCTATGAACTCATCAAGTTAAAGTTGACAGGCGCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
Db 324 CCTGTAGTCTCCTCCTCATCATG 347
```

```
RESULT 14
US-09-907-969-5
; Sequence 5, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Rling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Hall, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907.969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 536
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-969-5
```

```
Alignment Scores:
Pred. No.: 1.22e-47 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 10 Gaps: 0
```

US-09-720-934-2_copy_15_102 (1-88) x US-09-907-969-5 (1-531)

```
QY 1 TTPAlaIleThrValGlnGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db 84 TGGGCTATTACATCTGAAGAACGACTAAGCATGATAAACAAGTTGATTAACCTCAAACT 143
QY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnDhePheGlnSerGlyLeuPro 40
Db 144 TCAGAGGTTACATTAACAGGTGATCAAGCCCGTACTTTTCTTCCATCAGTCAAGTCTGCCG 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetLysnAspGlyArgMetAsp 60
Db 204 GCCCGGTTTACTTAAATATGGCCCTTATCAGATCTGAACAGATGGAGATGGAC 263
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysGlnGlyTyrGlnLeu 80
Db 264 CAGCAAGAGTTCTCTATAGCTATGAACTCATCAAGTTAAAGTTGACAGGCGCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
Db 324 CCTGTAGTCTCCTCCTCATCATG 347
```

RESULT 15

US-09-827-271-5

```
; Sequence 5, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C6
; CURRENT APPLICATION NUMBER: US/09/827.271
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-827-271-5
```

```
Alignment Scores:
Pred. No.: 1.22e-47 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 10 Gaps: 0
```

US-09-720-934-2_copy_15_102 (1-88) x US-09-827-271-5 (1-531)

```
QY 1 TTPAlaIleThrValGlnGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db 84 TGGGCTATTACATCTGAAGAACGACTAAGCATGATAAACAAGTTGATTAACCTCAAACT 143
QY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnDhePheGlnSerGlyLeuPro 40
Db 144 TCAGAGGTTACATTAACAGGTGATCAAGCCCGTACTTTTCTTCCATCAGTCAAGTCTGCCG 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetLysnAspGlyArgMetAsp 60
Db 204 GCCCGGTTTACTTAAATATGGCCCTTATCAGATCTGAACAGATGGAGATGGAC 263
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysGlnGlyTyrGlnLeu 80
Db 264 CAGCAAGAGTTCTCTATAGCTATGAACTCATCAAGTTAAAGTTGACAGGCGCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
Db 324 CCTGTAGTCTCCTCCTCATCATG 347
```

Search completed: July 1, 2004, 20:15:43
Job time : 275.977 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 11:41:07 ; Search time 271.625 Seconds

(without alignments)

1376.313 Million cell updates/sec

Title: US-09-720-934-2_COPY_15_102
Perfect score: 459
Sequence: 1 WAITVERAHDQGFHSLKP.....KLTKLKGQYQLSALPPVM 88

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 212409941 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xip
-Q=/cgm2_1/USPRO.spool.p/US097020934/runat.30062004.064539.13422/abp.query.fasta_1.1386
-DB=N_Geneseq.293an04 -OPMT=fastcap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR_MIN=0 -ALIGN=15
-MOB=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09720934.@CGN_1_1_1596.@runat.30062004.064539.13422 -NCPU=6 -ICPU=3
-NO_MMAP -LARGOUDERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.293an04.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002s.*
7: geneseqn2003as.*
8: geneseqn2003bs.*
9: geneseqn2003cs.*
10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	459	100.0	877	4	AAK93179 Human CDN
2	459	100.0	877	4	AAK91610 Human CDN
3	459	100.0	1676	4	AAK94611 Human SH3
4	459	100.0	2079	2	AAZ34573 Human SH3
5	459	100.0	2131	4	AAK94139 Human SH3
6	459	100.0	2131	4	AAH16578 Human SH3
7	459	100.0	3319	4	ABK43498 Human SH3
8	459	100.0	3466	4	AAI63825 Human SH3

9	459	100.0	5195	2	AAZ34572 Human SH3
10	459	100.0	5195	2	AAZ34570 Human SH3
11	459	100.0	5458	2	AAZ34571 Human SH3
12	453	98.7	7435	5	AAK84763 DNA encod
13	441	96.1	3723	3	AAZ39009 Mouse E
14	441	96.1	5082	3	AAZ39008 Mouse E
15	441	96.1	5144	3	AAZ39025 Mouse E
16	441	96.1	5738	3	AAZ39024 Mouse E
17	426	92.8	676	4	AAH08146 Human SH3
18	425.5	92.7	568	4	AAI63919 Human SH3
19	425.5	92.7	568	4	AAI63919 Human SH3
20	425.5	92.7	568	4	AAI63919 Human SH3
21	365	79.5	5828	6	AAI47247 Allergic
22	365	79.5	6103	4	AAI52332 Human SH3
23	359	78.2	480	3	AAI69750 Human SH3
24	359	78.2	480	6	ABN72644 Ovarian c
25	359	78.2	480	8	ADA08809 Human ova
26	359	78.2	531	3	AAI69695 Human ova
27	359	78.2	531	6	ABN72589 Ovarian c
28	359	78.2	531	8	ADA08754 Human ova
29	359	78.2	2017	3	AAI69762 Human ova
30	359	78.2	2017	6	ABN72656 Ovarian c
31	359	78.2	2017	8	ADA08821 Human ova
32	359	78.2	3593	3	AAZ39011 Mouse E
33	359	78.2	4625	3	AAZ39010 Mouse E
34	359	78.2	4975	3	AAZ39027 Mouse E
35	359	78.2	6014	3	AAZ39026 Mouse E
36	351	76.5	462	8	ACH34516 Human end
37	339	73.9	967	4	AAI63832 Human SH3
38	339	73.9	967	4	AAI63832 Human SH3
39	339	73.9	967	4	AAI63832 Human SH3
40	291	63.4	4447	4	AAI63821 DNA encod
41	262	57.1	3981	4	AAI63821 DNA encod
42	240	52.3	7225	4	ABL01994 Human SH3
43	212	46.2	292	4	AAI22800 Human SH3
44	212	46.2	292	4	ABA67887 Human SH3
45	212	46.2	292	4	AAI48103 Human SH3

ALIGNMENTS

RESULT 1
AAK93179
ID AAK93179 standard; CDN: 877 BP.
AC AAK93179;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human CDNA clone representative sequence, SEQ ID NO: 1639.
XX
KW Human, full length CDNA; CDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
EN EPI130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
XX
PR 11-JAN-2000; 2000JP-00118774.
XX
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length CDNA clones and their use

PT in genetic manipulation.

XX Example 11: SEQ ID NO 1639; 1380bp + Sequence Listing; English.

CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5' and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesized by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence was used as the representative sequence
 CC from a human clone which was used in homology searches to identify the
 CC clone. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in CD-ROM format directly from
 CC EPO

CC Sequence 877 BP; 208 A; 228 C; 223 G; 213 T; 0 U; 5 Other;

Alignment Scores:
 Pred. No.: 2,446-58 Length: 877
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AAK93179 (1-877)

QY 1 TTPAIAIETThVaIGluGluArgAlaYshHsAspGInGlnPheHsSerLeuYsPro 20
 DB 247 TGGGCGATTAACGTGAGAGAAAGAGGAGCATGATCAGCATTCATGTTAAAGCCA 306
 QY 21 ILESerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
 DB 307 ATATCTGATTCATTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTACT 366
 QY 41 GlnProValIleuAlaGlnIleTTPAlaIleuAlaAspMetAsnAsnAspGlyArgMetLeu 60
 DB 367 CAACCTGTTTTCAGACAGATATGGGCACTGACATGATTAATGATGGAAGATGAT 426
 QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
 DB 427 CAAGTGGAGTTTTCATAGCTATGAAACTTATCAACTCAAGACTCAAGATATACGCTA 486
 QY 81 ProSerAlaIleuProProValMet 88
 DB 487 CCCTCTGCACCTCCCTGTCATG 510

RESULT 2
 AAK91610
 ID AAK91610 standard; cDNA; 877 BP.

XX AAK91610;
 AC 06-NOV-2001 (first entry)
 DT
 XX Human CDNA 5'-end sequence, SEQ ID NO: 70.
 DE Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
 KM
 XX Homo sapiens.
 OS
 XX EPI130094-A2.
 PN
 XX 07-JUL-2000; 2000EP-00114089.
 PF
 XX 08-JUL-1999; 99JP-001944486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI: 2001-52425/58.

PT 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.

PS Claim 2; SEQ ID NO 70; 1380bp + Sequence Listing; English.

CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5' and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesized by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is the nucleotide sequence of the 5'-end of
 CC a cDNA provided in the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in CD-
 CC ROM format directly from EPO

CC Sequence 877 BP; 208 A; 228 C; 223 G; 213 T; 0 U; 5 Other;

Alignment Scores:
 Pred. No.: 2,446-58 Length: 877
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AAK91610 (1-877)

QY 1 TTPAIAIETThVaIGluGluArgAlaYshHsAspGInGlnPheHsSerLeuYsPro 20
 DB 247 TGGGCGATTAACGTGAGAGAAAGAGGAGCATGATCAGCATTCATGTTAAAGCCA 306
 QY 21 ILESerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
 DB 307 ATATCTGATTCATTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTACT 366
 QY 41 GlnProValIleuAlaGlnIleTTPAlaIleuAlaAspMetAsnAsnAspGlyArgMetLeu 60
 DB 367 CAACCTGTTTTCAGACAGATATGGGCACTGACATGATTAATGATGGAAGATGAT 426
 QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
 DB 427 CAAGTGGAGTTTTCATAGCTATGAAACTTATCAACTCAAGACTCAAGATATACGCTA 486
 QY 81 ProSerAlaIleuProProValMet 88
 DB 487 CCCTCTGCACCTCCCTGTCATG 510

RESULT 3
 AAK94611
 ID AAK94611 standard; cDNA; 1676 BP.

XX AAK94611;
 AC 07-NOV-2001 (first entry)
 DT
 XX Human CDNA, SEQ ID NO: 3565.
 DE Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
 KM
 XX Homo sapiens.
 OS
 XX EPI130094-A2.
 PN

```

XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-00114089.
PF
XX
XX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
DR
XX P-FSDB; AAM93676.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
in genetic manipulation.
XX
XX Disclosure; SEQ ID NO 3565; 1380bp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesizing the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesised by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a human cDNA provided in the
XX CC specification. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in CD-ROM format directly
XX from EPO
XX
SQ Sequence 1676 BP; 485 A; 381 C; 458 G; 352 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 5.91e-58 Length: 1676
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_15_102 (1-88) x AA934611 (1-1676)
QY 1 TnpalalrhrValGlugluArgAlaYshIsaspGInginPheHIsSerLeuYsPro 20
Db 306 TGGGCGCATTAAGTGAAGAAAGCCAGAGCATGATCGAGTTCATAGTTTAAAGCCA 365
QY 21 TlSerGlyPheIlethngIyAspGlnAlaArgAsnpePheGlnSerGlyLeuPro 40
Db 366 ATATCTGGAATTCATTACTGCGTGAATCAAGCTAGAACTTTTTCATCATCGGTTTACT 425
QY 41 GlnproValleuAlaGlnIleTnpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db 426 CCACTGTTTATGACACAGATATGCGCACTACCTACATGAATATGATGGAAGATGAT 485
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuYsLeuGlnGlyTYrGlnLeu 80
Db 486 CAACTGAGATTTCATACATAGCATAGAACTTTCAACTGAAGCTACAAAGATATCAGCTA 545
QY 81 ProSerAlaLeuProProValMet 88
Db 546 CCTCTGCACTTCCCCGTCATG 569
RESULT 4
AAZ34573
ID AAZ34573 standard; cDNA; 2079 BP.
XX
XX AC AAZ34573;
XX

```

```

DT 01-FEB-2000 (first entry)
XX
XX Human SH3D1A cDNA clone 5.
DE
XX
XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 136..2079
FT CDS /*tag= a
FT
XX
XX WC9953062-A2.
XX
XX 21-OCT-1999.
XX
XX 16-APR-1999; 99WO-US008371.
XX
XX 16-APR-1998; 98US-0082007P.
XX
XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
XX Korenberg JR, Chen X;
XX
XX WPI; 1999-633829/54.
XX
XX P-FSDB; AAY32157.
XX
XX Nucleic acid from the human SH3D1A gene and its products, useful for the
XX diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
XX Claim 2; Fig 12; 99JP; English.
XX
XX This is the nucleotide sequence of a cDNA clone, termed clone 5,
XX corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
XX contributes to the development of platelets and the pathogenesis of
XX leukaemias, both in general and in particular those involving the
XX megakaryocytic lineage. The SH3D1A gene maps to the small candidate
XX region for low platelets on chromosome 21. Sequencing of 5 different
XX sizes of cDNA clone from foetal brain (see AAZ34570-74) suggests that at
XX least 3 isoforms exist. The invention provides methods for the diagnosis
XX and treatment of megakaryocytic abnormality, myeloproliferative disorder,
XX platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
XX association of gains in chromosome 21 with leukaemias, neural
XX abnormalities, dysfunctions and disorders including brain malformations
XX and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
XX and colpocephaly. Methods are also provided for: suppressing cells unable
XX to regulate themselves; screening for a somatic alteration in the SH3D1A
XX gene; monitoring the progress and adequacy of a treatment; monitoring
XX tumour risk progress or megakaryocytic abnormality, myeloproliferative
XX disorder, haematopoietic disorder, platelet disorder or leukaemia; and
XX treatment of a subject (including a prenatal subject) having
XX megakaryocytic abnormality, myeloproliferative disorder, platelet
XX disorder, leukaemia or neural disorder using a nucleic acid that
XX expresses SH3D1A or its antisense nucleic acid
XX
SQ Sequence 2079 BP; 700 A; 451 C; 495 G; 433 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 7.93e-58 Length: 2079
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_15_102 (1-88) x AAZ34573 (1-2079)
QY 1 TnpalalrhrValGlugluArgAlaYshIsaspGInginPheHIsSerLeuYsPro 20

```

```

Db      178 TGGGCCATACCTGAGAGAAAGAGGAGCATGATCAGACATTCATGTTAAAGCA 237
Qy      21 11eSerGlyPhe11eThrGlyaspGln1aAargAsnPhenPheGlnSerGlyLeuPro 40
Db      238 AATCTGGAATTCATCTGATGATCAAGCTAAGAACTTTTTCATCTGGGTTACCT 297
Qy      41 GlnProValLeuAlaGln11eTrrPa1aLeuAlaaspMetAsnAsnAspGlyArgMetAsp 60
Db      238 CAACCTGTTTTCAGACAGATATGGGCACTAGCTGACATGATATGATGAAAGATGAT 357
Qy      61 GlnValGluPheSer11eAlaMetLysLeu11eLysLeuGlnGlyArgLeu 80
Db      358 CAAGTGAAGTTTTCATAGCTATGAAGCTTATCAAGCTACAAAGATATACGCTA 417
Qy      81 ProSerAlaLeuProProValMet 88
Db      418 CCTCTGCACTTCCCCCTGTCTCATG 441

```

RESULT 5

AAK94139 standard; cDNA; 2131 BP.

AAK94139;
06-NOV-2001 (first entry)

Human full-length cDNA, SEQ ID NO: 2646.

Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

Homo sapiens.

EP1130094-A2.

05-SEP-2001.

07-JUL-2000; 2000BP-00114089.

08-JUL-1999; 99JP-00194486.

11-JAN-2000; 2000JP-00118774.

02-MAY-2000; 2000JP-00183765.

(HELI-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

WPI; 2001-524255/58.

P-PsDB; AAM93229.

830 Primers useful for synthesizing full length cDNA clones and their use

in genetic manipulation.

Claim 8; SEQ ID NO 2646; 1380bp + Sequence Listing; English.

The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5' and 3'-ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesized by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a full length human cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO

Sequence 2131 BP; 694 A; 469 C; 526 G; 442 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.2e-58

Length: 2131

Score: 459.00 Matches: 88
Percent Similarity: 100.00%
Percent Identity: 100.00%
Query Match: 100.00%
DB: 4
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AAK94139 (1-2131)

```

Qy      1 TrrPa1a11eThValG1uG1uArgAla1ySH1sAspGlnGlnPheHisSerLeuLysPro 20
Db      247 TGGGCCATACCTGAGAGAAAGAGGAGCATGATCAGACATTCATGTTAAAGCA 306
Qy      21 11eSerGlyPhe11eThrGlyaspGln1aAargAsnPhenPheGlnSerGlyLeuPro 40
Db      307 AATCTGGAATTCATCTGATGATCAAGCTAAGAACTTTTTCATCTGGGTTACCT 366
Qy      41 GlnProValLeuAlaGln11eTrrPa1aLeuAlaaspMetAsnAsnAspGlyArgMetAsp 60
Db      367 CAACCTGTTTTCAGACAGATATGGGCACTAGCTGACATGATATGATGAAAGATGAT 426
Qy      61 GlnValGluPheSer11eAlaMetLysLeu11eLysLeuGlnGlyArgLeu 80
Db      427 CAAGTGAAGTTTTCATAGCTATGAAGCTTATCAAGCTACAAAGATATACGCTA 486
Qy      81 ProSerAlaLeuProProValMet 88
Db      487 CCTCTGCACTTCCCCCTGTCTCATG 510

```

RESULT 6

AAH16578 standard; cDNA; 2131 BP.

AAH16578;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:15658.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000BP-00116126.

29-JUL-1999; 99JP-00248036.

27-AUG-1999; 99JP-00300253.

11-JAN-2000; 2000JP-00118776.

02-MAY-2000; 2000JP-00183765.

09-JUN-2000; 2000JP-00241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 15658; 2537bp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AA892446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX
SQ Sequence 2131 BP; 691 A; 448 C; 572 G; 420 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.2e-58 Length: 2131
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AAH16578 (1-2131)

QY 1 TrrAlalethbValGlugluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db TGGCCCAATACGTAGAGAGAAAGCCAGCATATACGACGTTCATAGTTTAAACCA 448
QY 21 TleserGIpheelrhrGIyAspGlnAlaArgasnPhepheGlnSerGIyLysPro 40
Db ATATCTGGATTTCATTACGTGATCAAGCTAGAACTTTTTCATTCTGGGTTTACT 508
QY 41 GlnProValIleuAlaGlnIleTrrAlaIleuAlaAspMetCnaanaaspGIyArgMetAsp 60
Db CAACCTGTTTATAGACAGATATGGCGCATGCTGACATGATATATGATGGAAGATGCAT 568
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnIyTyrGlnLeu 80
Db CAACTGAGATTTCATAGCTATGAACTTATCAAACTGAACTCAAGATATCAGCTA 628
QY 81 ProSerAlaIleuProProValMet 88
Db CCCCTGCACTCCCTCCGTGATG 652

RESULT 7
ABK43498
ID ABK43498 standard; cDNA; 3319 BP.
XX
XX ABK43498;
AC
XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding novel central nervous system protein #78.
XX
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.
XX
XX Homo sapiens.
XX

PN W0200155318-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001332.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234938P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.
XX 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
XX 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
XX 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
XX 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
XX 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
XX 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
XX 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
XX 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
XX 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
XX 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
XX 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
XX 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
XX 08-DEC-2000; 2000US-0251864P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
XX 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI

XX MPI: 2001-581633/65.
DR P-PSDB; AAU87168.
XX
PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
PR treating or ameliorating medical conditions and used as food additives or
PT preservatives.
XX
PS Claim 1; SEQ ID NO 88; 837pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (II) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infection. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
Alignment Scores:
Pred. No.: 1.5e-57 Length: 3319
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_15_102 (1-88) x ABK43498 (1-3319)
QY 1 TTPAlaIleThnValGluGluArgAlaIysHISAspGlnGlnPheHISerLeuIysPro 20
Db 201 TGGGCCATTAAGTGTAGAGGAAAGACGAGGATGATGACCACTTCATATGTTAAAGCCA 260
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAenPhePheGlnserGlyLeuPro 40
Db 261 ATATCTGATTCATTAAGTGTATCATCAAGCTAAGAACTTTTTCATATCTGGTTACT 320
QY 41 GlnProValIleuAlaGlnIleThrPalaleuAlaaspMetAsnAspGlyArgMetAsp 60
Db 321 CAACCTGTTTGAACACATATAGGACCTAGCTGACATGAATAAATGATGAAAGATGAT 380
QY 61 GlnValGluPheSerIleAlaMetIysLeuIleIysLeuIysLeuGlnIysIleTyrGlnLeu 80
Db 381 CAAGTGAAGTTTTCATATGACTATGAATTAATCAATCAAGCTACAGAGATATACGCTA 440
QY 81 ProSerAlaLeuProProValMet 88
Db 441 CCCTCTGACATTCCTCCCTGTCTATG 464
RESULT 8
AA163825
ID AA163825 standard; cDNA, 3466 BP.
XX
XX AA163825;
AC
XX
DT 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 33.
DE

XX Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
KW neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antitumor; vulnery; anticonvulsant; antibacterial;
KW antiparastic; cardiac; gene therapy; cancer; immune disorder;
KW cardiovascular disorder; neurological disease; infection; human; ss.
XX Homo sapiens.
XX WO200155308-A2.
XX 02-AUG-2001.
PD 17-JAN-2001; 2001WO-US001309.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184654P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215115P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241866P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249248P.
PR 17-NOV-2000; 2000US-0249249P.
PR 17-NOV-2000; 2000US-0249250P.
PR 17-NOV-2000; 2000US-0249251P.
PR 17-NOV-2000; 2000US-0249252P.
PR 17-NOV-2000; 2000US-0249253P.
PR 17-NOV-2000; 2000US-0249254P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249258P.
PR 17-NOV-2000; 2000US-0249259P.
PR 17-NOV-2000; 2000US-0249260P.
PR 17-NOV-2000; 2000US-0249261P.
PR 17-NOV-2000; 2000US-0249262P.
PR 17-NOV-2000; 2000US-0249263P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249268P.
PR 17-NOV-2000; 2000US-0249269P.
PR 17-NOV-2000; 2000US-0249270P.
PR 17-NOV-2000; 2000US-0249271P.
PR 17-NOV-2000; 2000US-0249272P.
PR 17-NOV-2000; 2000US-0249273P.
PR 17-NOV-2000; 2000US-0249274P.
PR 17-NOV-2000; 2000US-0249275P.
PR 17-NOV-2000; 2000US-0249276P.
PR 17-NOV-2000; 2000US-0249277P.
PR 17-NOV-2000; 2000US-0249278P.
PR 17-NOV-2000; 2000US-0249279P.
PR 17-NOV-2000; 2000US-0249280P.
PR 17-NOV-2000; 2000US-0249281P.
PR 17-NOV-2000; 2000US-0249282P.
PR 17-NOV-2000; 2000US-0249283P.
PR 17-NOV-2000; 2000US-0249284P.
PR 17-NOV-2000; 2000US-0249285P.
PR 17-NOV-2000; 2000US-0249286P.
PR 17-NOV-2000; 2000US-0249287P.
PR 17-NOV-2000; 2000US-0249288P.
PR 17-NOV-2000; 2000US-0249289P.
PR 17-NOV-2000; 2000US-0249290P.
PR 17-NOV-2000; 2000US-0249291P.
PR 17-NOV-2000; 2000US-0249292P.
PR 17-NOV-2000; 2000US-0249293P.
PR 17-NOV-2000; 2000US-0249294P.
PR 17-NOV-2000; 2000US-0249295P.
PR 17-NOV-2000; 2000US-0249296P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.


```
US-09-720-934-2_COPY_15_102 (1-88) x AAZ34572 (1-5195)
QY 1 TTPAlalleThrValGluGluAArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 281 TGGGCCATTAACCTAGAGGAAAGCGAAGCATGATCAAGATTCATGTTTAAAGCCA 340
QY 21 TlSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
DB 341 ATATCTGGATTCACTACTGCTGATCAAGCTAAACTTTTTCATCTGGGTTACCT 400
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 401 CAACCTGTTTATAGACAGATATGGGCACTAGCTACATGAATATGATGGAAGATGGAT 460
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 461 CAAGTGGAGTTTTCATAGCTATGAACTTATCAAACTGAAGCTACAAGATATCAAGCTA 520
QY 81 ProSerAlaLeuProProValMet 88
DB 521 CCTCTGCACCTCCCTGTCATG 544
RESULT 10
AAZ34570
ID AAZ34570 standard; cDNA; 5199 BP.
AC AAZ34570;
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA.
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW hematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 208..3642
FT /*tag= a
XX
PN WO9953062-A2.
XX
XX 21-OCT-1999.
XX
PF 16-APR-1999; 99WC-US008371.
XX
PR 16-APR-1998; 98US-0082007P.
XX
PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
PI Korenberg JR, Chen X;
XX
DR WPI, 1999-633829/54.
XX
DR P-PSDB; AAY32154.
XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
PT diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
PS Claim 2, Fig 5, 9pp; English.
XX
XX This is the nucleotide sequence of full-length cDNA corresponding to a
XX novel human SH3 gene, termed the SH3D1A gene, that contributes to the
XX development of platelets and the pathogenesis of leukemias, both in
XX general and in particular those involving the megakaryocytic lineage. The
XX SH3D1A gene maps to the small candidate region for low platelets on
XX chromosome 21. Sequencing of 5 different sizes of cDNA clone (see
XX AAZ34570-74) suggests that at least 3 isoforms exist. The invention
XX provides methods for the diagnosis and treatment of megakaryocytic
XX abnormality, myeloproliferative disorder, platelet disorder, acute
```

```
CC leukaemia, neural disorders, thrombocytopenia, platelet disorder on
CC chromosome 21, low platelets in deletion for 21, association of gains in
CC chromosome 21 with leukemias, neural abnormalities, dysfunctions and
CC disorders including brain malformations and corresponding cognitive
CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
CC also provided for: suppressing cells unable to regulate themselves;
CC screening for a somatic alteration in the SH3D1A gene; monitoring the
CC progress and adequacy of a treatment; monitoring tumour risk progress or
CC megakaryocytic abnormality, myeloproliferative disorder, hematopoietic
CC disorder, platelet disorder or leukaemia; and treatment of a subject
CC (including a prenatal subject) having megakaryocytic abnormality,
CC myeloproliferative disorder, platelet disorder, leukaemia or neural
CC disorder using a nucleic acid that expresses SH3D1A or its antisense
CC nucleic acid
XX
SQ Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2.78e-57 Length: 5199
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
XX
US-09-720-934-2_COPY_15_102 (1-88) x AAZ34570 (1-5199)
QY 1 TTPAlalleThrValGluGluAArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 250 TGGGCCATTAACCTAGAGGAAAGCGAAGCATGATCAAGATTCATGTTTAAAGCCA 309
QY 21 TlSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
DB 310 ATATCTGGATTCACTACTGCTGATCAAGCTAAACTTTTTCATCTGGGTTACCT 369
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 370 CAACCTGTTTATAGACAGATATGGGCACTAGCTACATGAATATGATGGAAGATGGAT 429
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 430 CAAGTGGAGTTTTCATAGCTATGAACTTATCAAACTGAAGCTACAAGATATCAAGCTA 489
QY 81 ProSerAlaLeuProProValMet 88
DB 490 CCTCTGCACCTCCCTGTCATG 513
RESULT 11
AAZ34571
ID AAZ34571 standard; cDNA; 5458 BP.
XX
AC AAZ34571;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA clone 21.
XX
XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW hematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
XX Homo sapiens.
XX
OS
XX
FH Key Location/Qualifiers
FT CDS 267..3929
FT /*tag= a
XX
PN WO9953062-A2.
XX
XX 21-OCT-1999.
XX
```

PF 16-APR-1999; 99WO-US008371.
XX
PR 16-APR-1998; 98US-0082007P.
XX
XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
PI Korenberg JR, Chen X;
XX
DR MPI; 1999-633829/54.
XX P-FSDB; AAY32155.
PT
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
PT diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
PS Claim 2; Fig 8; 99pp; English.

CC This is the nucleotide sequence of full-length cDNA (clone 21)
CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
CC contributes to the development of platelets and the pathogenesis of
CC leukemias, both in general and in particular those involving the
CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
CC region for low platelets on chromosome 21. Sequencing of 5 different
CC sizes of cDNA clone from foetal brain (see AAZ34570-74) suggests that at
CC least 3 isoforms exist. The invention provides methods for the diagnosis
CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
CC platelet disorder on chromosome 21, low platelets in deletion for 21,
CC association of genes in chromosome 21 with leukaemias, neural
CC abnormalities, dysfunction and disorders including brain malformations
CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
CC and colpocephaly. Methods are also provided for: suppressing cells unable
CC to regulate themselves; screening for a somatic alteration in the SH3D1A
CC gene; monitoring the progress and adequacy of a treatment; monitoring
CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
CC treatment of a subject (including a prenatal subject) having
CC megakaryocytic abnormality, myeloproliferative disorder, platelet
CC disorder, leukaemia or neural disorder using a nucleic acid that
CC expresses SH3D1A or its antisense nucleic acid

XX
SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,976-57 Length: 5458
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AAZ34571 (1-5458)

QY 1 TTPAAlEtHrVAlGluGluAlaYshHSaRpGInGInPheHISerLeuYsPro 20
Db 309 TGGGCAATACCTAGTAGAGAAAGACGACATGATCGACATTCATATGTTAAAGCCA 368
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
Db 369 ATATCTGATTCATCTACTGCGTATCAAGCTAGAAACCTTTTTTTCATCTGGGTTACT 428
QY 41 GlnProValLeuAlaGlnIleThrAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db 429 CAACCTGTTTAGCAGACATATGGGCACTAGCTGACATGAATATGATGGAAGATGAT 488
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuValGluGlyTyrGlnLeu 80
Db 489 CAAGTGGAGTTTTCATAGCTATGAACTTATCAACTTAACTGAAGATATCAAGTAA 548
QY 81 ProSerAlaLeuProProValMet 88
Db 549 CCTCTGACATTCCTCCCTCTCATNG 572

RESULT 12

AAS84763
ID AAS84763 standard; cDNA; 7435 BP.
XX
AC AAS84763;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #20567.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM Food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR MPI; 2001-639362/73.
XX P-FSDB; ABG20576.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 1; SEQ ID NO 20567; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,566-56 Length: 7435
Score: 453.00 Matches: 87
Percent Similarity: 98.86% Conservative: 0
Best Local Similarity: 98.86% Mismatches: 1
Query Match: 98.69% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AAS84763 (1-7435)

QY 1 TTPAAlEtHrVAlGluGluAlaYshHSaRpGInGInPheHISerLeuYsPro 20
|||||

```
Db 305 TGGGCCAATGCTGAGAGGAAGGAGAGCATGATCAGCAGCTCCATGTTTAAAGCCA 364
QY 21 ILESerGlyPheIleThrgIYAAspGlnAlaArgSnpPhePheGlnSerGlyLeuPro 40
Db 365 ATATCTGATTCATCTACTGCGTGAATCAAGCTAAGAACTTTTTCATCTGGGTTACT 424
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetSnaAsnAspGlyArgMetAsp 60
Db 425 CAACCTGTTTATGACAGATATGGGCACTGACATGATATATGATGGAAGATGAT 484
QY 61 GlnValGlnPheSerIleAlaMetIleLeuIleLysLeuLysLeuGlnGlyTYrGlnLeu 80
Db 485 CAACTGGAATTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAAAGATATCAGCTA 544
QY 81 ProSerAlaLeuProProValMet 88
Db 545 CCTCTGCACTTCCCTGTCATG 568

RESULT 13
AAZ39009
ID AAZ39009 standard; cDNA; 3723 BP.
XX
AC AAZ39009;
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse Esei coding sequence.
XX
KM Mouse; murine; Esei; Esee2; endocytosis; vesicular trafficking;
KM regulation; actin cytoskeleton; detection; cancer; infection;
KM EH-domain and SH3-domain regulator of endocytosis; anticancer;
KM antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
PN WO9955728-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
DR WPI: 2000-052802/04.
DR P-PSDB; AAY57444.
XX
PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS Claim 6; Page 40-42; 99pp; English.
XX
CC The present sequence encodes mouse Esei. The present invention
CC specifically describes mammalian Esei and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp15 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Esei is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
```

```
CC cultures, by forming an Ese-Esp15 complex, then binding dynamn to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 8.52e-55 Length: 3723
Score: 441.00 Matches: 85
Percent Similarity: 97.73% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 2
Query Match: 96.08% Indels: 0
DB: 3 Gaps: 0
US-09-720-934-2_copy_15_102 (1-88) x AAZ39009 (1-3723)
QY 1 TTPAlaIleThrValGluGluArgAlaLysAspGlnGlnPheHisSerLeuLysPro 20
Db 43 TGGGCCAATGCTGAGAGGAAGGAGGAGCATGATGACGAGATTCCTTAGCTGAGCCG 102
QY 21 ILESerGlyPheIleThrgIYAAspGlnAlaArgSnpPhePheGlnSerGlyLeuPro 40
Db 103 ATAGCGGATTTATTAATGATGATCAAGCAGGAACTTTTTCATCTGGGTTACT 162
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetSnaAsnAspGlyArgMetAsp 60
Db 163 CAGCTGCTTATGACAAATATGGCCCTGCGCATGATTAACATGAGATGATGAT 222
QY 61 GlnValGlnPheSerIleAlaMetIleLeuIleLysLeuLysLeuGlnGlyTYrGlnLeu 80
Db 223 CAACTGGAATTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAAAGATATCAGCTC 282
QY 81 ProSerAlaLeuProProValMet 88
Db 283 CCTCCACACTTCCCTGTCATG 306

RESULT 14
AAZ39008
ID AAZ39008 standard; cDNA; 5082 BP.
XX
AC AAZ39008;
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse Esei full length cDNA sequence.
XX
KM Mouse; murine; Esei; Esee2; endocytosis; vesicular trafficking;
KM regulation; actin cytoskeleton; detection; cancer; infection;
KM EH-domain and SH3-domain regulator of endocytosis; anticancer;
KM antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
PN WO9955728-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
DR WPI: 2000-052802/04.
DR P-PSDB; AAY57444.
XX
PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
```

PS Claim 6; Page 38-40; 99pp; English.

XX The present sequence encodes mduse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp1 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1 complex, then binding dynam to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signaling, tissue development or synaptic transmission
CC XX

Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.3e-54	Length:	5082
Score:	441.00	Matches:	85
Percent Similarity:	97.73%	Conservative:	1
Best Local Similarity:	96.59%	Mismatches:	2
Query Match:	96.08%	Indels:	0
DB:	3	Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) x AAZ39008 (1-5082)

QY	1	TrypAlaIeThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuIysPro	20
DB	301	TGGGCGATACACTGTGGAGGAAAGGCCCAAGCATGACGACATCTTAACTGAAAGCCG	360
QY	21	IlleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro	40
DB	361	ATAGCGGATTTATTAATCTGTCATCAAGCGAGAACTTTTTCCAATCTGGTTACT	420
QY	41	GlnProValIleuAlaGlnIleThrAlaLeuAlaAspMetAsnAspGlyArgMetAsp	60
DB	421	CAGCTGTCTTAGCAACAATATGGCGCTAGCGGACATGATTAACATGGAAGATGAA	480
QY	61	GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu	80
DB	481	CAAGTGAATTTTCCATAGCCATGAAGCTTATCAAACTGAAGCTTCAAGATATCAGCTC	540
QY	81	ProSerAlaLeuProProValMet	88
DB	541	CCCTCCACACTTCCCTCGTCATG	564

RESULT 15

AAZ39025

AAZ39025 standard; cDNA, 5144 BP.

AAZ39025;

28-FEB-2000 (first entry)

Mouse Ese1 coding sequence.

Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;

regulation; actin cytoskeleton; detection; cancer; infection;

EH-domain and SH3-domain regulator of endocytosis; anticancer;

antiproliferative; antiviral; ss.

Mus sp.

PN WO955728-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.

XX 27-APR-1999; 98CA-02230201.

XX 05-FEB-1999; 99US-0118739P.

XX (HSCR-) HSC RES & DEV LP.

PI Egan SE, Wang W, Sengar A;

XX MPI: 2000-052802/04.

XX P-PSDB; AA57443.

PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of

XX endocytosis, used e.g. for treating cancer or preventing viral infection.

XX Claim 6; Page 59-62; 99pp; English.

XX The present invention specifically describes mammalian Ese1 and 2
CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
CC regulator of endocytosis). (I) are involved in regulation of clathrin-
CC mediated endocytosis (as a complex with Esp1 protein), vesicular
CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
CC sequences antisense to the (I) polynucleotide; agents that downregulate
CC expression of Ese genes or antagonists of an Ese binding partner are used
CC to treat diseases associated with undesirable endocytosis and resulting
CC changes in cellular function. Particularly overexpression of Ese1 is used
CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
CC administration of (I) is used to promote endocytosis of selected cells.
CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
CC cells that can be stimulated to proliferate by a growth factor receptor;
CC and similar compounds (also inactive Ese mutants) can be used to prevent
CC viral infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1 complex, then binding dynam to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signaling, tissue development or synaptic transmission. The
CC present sequence represents mouse Ese1 coding sequence
CC XX

Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.33e-54	Length:	5144
Score:	441.00	Matches:	85
Percent Similarity:	97.73%	Conservative:	1
Best Local Similarity:	96.59%	Mismatches:	2
Query Match:	96.08%	Indels:	0
DB:	3	Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) x AAZ39025 (1-5144)

QY	1	TrypAlaIeThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuIysPro	20
DB	43	TGGGCGATACACTGTGGAGGAAAGGCCCAAGCATGACGACATCTTAACTGAAAGCCG	102
QY	21	IlleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro	40
DB	103	ATAGCGGATTTATTAATCTGTCATCAAGCGAGAACTTTTTCCAATCTGGTTACT	162
QY	41	GlnProValIleuAlaGlnIleThrAlaLeuAlaAspMetAsnAspGlyArgMetAsp	60
DB	163	CAGCTGTCTTAGCAACAATATGGCGCTAGCGGACATGATTAACATGGAAGATGAA	222
QY	61	GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu	80
DB	223	CAAGTGAATTTTCCATAGCCATGAAGCTTATCAAACTGAAGCTTCAAGATATCAGCTC	282
QY	81	ProSerAlaLeuProProValMet	88

, Fri Jul 2 10:23:36 2004

us-09-720-934-2_copy_15_102.rng

Page 13

Db 283 CCTTCACACTCCCCCTGTCATG 306

Search completed: July 1, 2004, 13:31:27
Job time : 278.625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 2189.79 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 WAHPQSSRLKYROLFNSHDK.....SGQPLPPVLPPEYIPSPFR 96

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=/cgml1/USPTO.spool.p/US09720934/runat.30062004.064539.13430/app.query.fasta_1.1386
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NOR=exc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09720934.@CGML1.1.7509.@runat.30062004.064539.13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-FEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgt_mus:*
34: em_hgt_pin:*
35: em_hgt_rod:*
36: em_hgt_mam:*
37: em_hgt_vrt:*
38: em_sy:*
39: em_hgt_hum:*
40: em_hgt_mus:*
41: em_hgt_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	501	100.0	1299	9 BC039036	BC039036 Homo sapi
2	501	100.0	1676	6 BD127640	BD127640 Primer fo
3	501	100.0	1676	9 AK075290	AK075290 Homo sapi
4	501	100.0	2079	6 BD205036	BD205036 Isolated
5	501	100.0	2131	6 AX880753	AX880753 Sequence
6	501	100.0	2131	6 BD127168	BD127168 Primer fo
7	501	100.0	2131	6 BD158570	BD158570 Primer fo
8	501	100.0	2131	9 AK027846	AK027846 Homo sapi
9	501	100.0	2131	9 AK074554	AK074554 Homo sapi
10	501	100.0	2199	9 BC058925	BC058925 Homo sapi
11	501	100.0	3723	10 AF132478	AF132478 Mus muscu
12	501	100.0	3812	10 AF132672	AF132672 Rattus no
13	501	100.0	4025	10 AF127798	AF127798 Rattus no
14	501	100.0	4321	9 HSM806384	HSM806384 Homo sapi
15	501	100.0	5145	10 AF132481	AF132481 Mus muscu
16	501	100.0	5195	6 BD205035	BD205035 Isolated
17	501	100.0	5199	6 BD205033	BD205033 Isolated
18	501	100.0	5287	9 AF064243	AF064243 Homo sapi
19	501	100.0	5381	9 AF114488	AF114488 Homo sapi
20	501	100.0	5458	6 BD205034	BD205034 Isolated
21	501	100.0	6439	9 AF114487	AF114487 Homo sapi
22	501	100.0	7247	9 AF064244	AF064244 Homo sapi
23	474	94.6	7105	5 AF032118	AF032118 Xenopus l
24	359.5	71.8	136937	2 AC139627	AC139627 Takifugu
25	349	69.7	2017	6 BD264687	BD264687 Compositi
26	349	69.7	2017	6 AR238090	AR238090 Sequence
27	349	69.7	2017	6 AR257631	AR257631 Sequence
28	349	69.7	2017	6 AR283677	AR283677 Sequence
29	349	69.7	2017	6 AX366305	AX366305 Sequence
30	349	69.7	3594	10 AF132480	AF132480 Mus muscu
31	349	69.7	4977	10 AF132479	AF132479 Mus muscu
32	349	69.7	5804	10 AK122480	AK122480 Mus muscu
33	346	69.1	1749	9 BC020921	BC020921 Homo sapi
34	346	69.1	2085	9 BC038963	BC038963 Homo sapi
35	345	68.9	4447	6 AX428899	AX428899 Sequence
36	345	68.9	4557	9 AF182199	AF182199 Homo sapi
37	345	68.9	5828	6 BD167848	BD167848 Method fo
38	345	68.9	5828	9 AF248540	AF248540 Homo sapi
39	345	68.9	5938	9 AB033082	AB033082 Homo sapi
40	345	68.9	6091	9 AF182198	AF182198 Homo sapi
41	292	58.3	3750	3 AF054612	AF054612 Drosophi
42	292	58.3	3999	3 AF053957	AF053957 Drosophi
43	292	58.3	127035	3 AC006574	AC006574 Drosophi
44	292	58.3	160282	3 AC009253	AC009253 Drosophi
45	292	58.3	273995	3 AF003669	AF003669 Drosophi

RESULT 1

ALIGNMENTS

BC039036
LOCUS BC039036 1299 bp mRNA linear PRI 24-DEC-2002
DEFINITION Homo sapiens, similar to intersectin 1 (SH3 domain protein), clone IMAGE:5784342, mRNA, partial cds.
ACCESSION BC039036
VERSION BC039036.1 GI:24659505
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1299)
TITLE Strausberg, R.
JOURNAL Submitted (01-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nih.gov
Aklter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, D.C., McQuell, J., Pearson, R., Stantipop, S., Thomas, P.D., Touchman, D.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Weherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAX Plate: 82 Row: 1 Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796.
FEATURES
source
1..1299
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5784342"
/tissue_type="uterus, leiomyosarcoma"
/clone_lib="NTH MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
264..>1299
/codon_start=1
/product="Similar to intersectin 1 (SH3 domain protein)"
/protein_id="AAH39036.1"
/db_xref="GI:24711335"
/translation="MAQPTTPGSGSLDIWATVERARHDOQFSLKISGFTTDDA
RNFFQSGLPQVLAQIWAALMNDGMDQVEFIANKLILKLGQSLBSALPVM
KOQVAVISAPAFGAGISAMPPLTAVAPVPMGS.PVVGMSPTLVSVPTAAPLAN
GAPVITQPLPAFAHATLTKSSSFRSGSGLTKLOKASPFVASVPPVAEVA
OSSRLKYQLFNSHKTKTSGHLTGQARTIMQSSLPQALASIMNLSIDDDGLTA
EEFLIAMTLDIVAMSGCLPVLPPPIPPRRRVRSSGSGISVISTVDQRLPEPV
LEDEQOQKKKK"

DB: 9 Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x BC039036 (1-1299)
QY 1 TTPALAVLPROGINSERSEARGLNYSRYARGLNLEUPEASNSERHISAPLYS 20
DB 906 TGGGCTGTTCCCTCACTATCAAGACTGAAATACAGGCAATTATTCATGTCATGACAAA 965
QY 21 TThMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 966 ACTATGAGTGGACACTTACAGTCCCAAGCAAGCAACATTCCTTATGACGTCAGATTAA 1025
QY 41 ProGlnAlaGlnLeuAlaSerIleTrrpAsnLeuSerAspIleAspGlnAspGlySerLeu 60
DB 1026 CCACAGGCTCAGCTGCTTCAATATGCAATCTTTCTGACATTGATCCAGATGGAACCTT 1085
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1086 ACAGCAGAGGAATTATTCCTGGCAATGCACCTCATTTGATGTATGTCGGCCACCA 1145
QY 81 LeuProProValLeuProProGlnTyrIleProProSerPheArgArg 96
DB 1146 CTGCCACCTGCTCTGCTCCACGAATACATTCACCTCTTTTAAAGA 1193
RESULT 2
LOCUS BD127640 1676 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127640
VERSION BD127640.1 GI:23222585
KEYWORDS JP 2002017375-A/3071.
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1676)
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 3071 22-JAN-2002;
COMMENT
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/3071
PD 22-JAN-2002
PI 07-JUN-2000 JP 2000253172
PT TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC
10,
PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N5/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
source
FT CDS Location/Qualifiers
1..1676
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ALIGNMENT SCORES
Pred. No.: 1.6e-50 Length: 1676
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x BD127640 (1-1676)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTYTArgGlnLeuPheAsnSerHisAspLys 20
DB 906 TGGGCTGTTCTCAGTCATCAAGACTGAAATACAGGCAATATTCACTCATGACAAA 965

QY 21 ThMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 966 ACTATGAGTGGACACTTAACAGGTCCTCCCAAGCAAGAACTATCTTATGACGTCAAGTTTA 1025

QY 41 ProGlnAlaGlnLeuAlaSerIleTPAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
DB 1026 CCACAGGCTCAGCTGCTGCTCAATATGAACTTTCTTGACATTCATCAAGATGAAAACTT 1085

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1086 ACAGCAGAGGAATTTATCTCGGCAATGACCTCATTTGATGATGCTCGGCCAACCA 1145

QY 81 LeuProProValLeuProProGlyTyrIleProProSerPheArgArg 96
DB 1146 CTGCCACCTGTCCTGCTCCAGAAATACATTCCACCTTTTGAAGA 1193

RESULT 3
AK075290 1676 bp mRNA linear PRI 03-SEP-2002
LOCUS Homo sapiens cDNA FLJ30809 fis, clone Y79PA1000778, highly similar
DEFINITION to Homo sapiens intersecin long form mRNA.
ACCESSION AK075290
VERSION AK075290.1 GI:22761283
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1
Iisogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hiro,Y., Saico,K., Yamamoto,J., Makamatsu,A., Nakamura,Y.,
Kojima,S., Nagahata,K., Masuno,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Kotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1676)
AUTHORS Iisogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2002) Takao Iisogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@nri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo; Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).

FEATURES
source
1..1676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Y79PA1000778"
/cell_line="Y79"
/cell_type="retinoblastoma"
/clone_lib="Y79A1"
/note="cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 1.6e-50 Length: 1676
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AK075290 (1-1676)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTYTArgGlnLeuPheAsnSerHisAspLys 20
DB 906 TGGGCTGTTCTCAGTCATCAAGACTGAAATACAGGCAATATTCACTCATGACAAA 965

QY 21 ThMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 966 ACTATGAGTGGACACTTAACAGGTCCTCCCAAGCAAGAACTATCTTATGACGTCAAGTTTA 1025

QY 41 ProGlnAlaGlnLeuAlaSerIleTPAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
DB 1026 CCACAGGCTCAGCTGCTGCTCAATATGAACTTTCTTGACATTCATCAAGATGAAAACTT 1085

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1086 ACAGCAGAGGAATTTATCTCGGCAATGACCTCATTTGATGATGCTCGGCCAACCA 1145

QY 81 LeuProProValLeuProProGlyTyrIleProProSerPheArgArg 96
DB 1146 CTGCCACCTGTCCTGCTCCAGAAATACATTCCACCTTTTGAAGA 1193

RESULT 4
BD205036 2079 bp DNA linear PAT 17-JUN-2003
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205036
VERSION BD205036.1 GI:33014806
KEYWORDS JP 2002511267-A/4.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 2079)
Korenberg,J.R. and Chen,X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
JOURNAL leukemia and utilization thereof
COMMENT Patent: JP 2002511267-A 4 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL
OS Homo sapiens (human)
PN JP 2002511267-A/4
PD 16-APR-2002
PR 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JUNIE R KORENBERG, XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

FEATURES
source
1..2079
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.02e-50 Length: 2079
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

Db	847	UGGGCTGTTCTCAGTCATCAAGACTGAAGTAATACAGCAATTAATTCATAGTCATGACAAA	908
Qy	21	ThrmSetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu	40
Db	907	ACTATGAGTGGACACTTAAACAGGTGCCCAAGCAAGAACTATCTTATCAGTCAAGTTTA	966
Qy	41	ProGlnAlaGlnLeuAlaSerIleThrPheLeuSerArgP1leAspGlnAspGlyLysIeu	60
Db	967	CCACAGGCTCAGCTGGCTTCATTAAGAAATCTTTCTGTGACATTGATCAAGATGGAACCTT	1026
Qy	61	ThraIaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro	80
Db	1027	ACAGAGAGAGGAATTTATCTCGGCATGACCTCATTTGATGTCATGTCCTGGCCACCA	1086
Qy	81	LeuProProValLeuProProGlnIleProGlnProSerPheArgArg	96
Db	1087	CTGCACCTGTCCTCCTCCAGAAATACATTCACCTCTTTTAGAAGA	1134
RESULT 7			
LOCUS	BD158570	2131 bp	DNA
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.		
ACCESSION	BD158570		
VERSION	BD158570.1	GI:278664328	
KEYWORDS	JP 2002191363-A/13413.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Ota,T., Isogai,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,J.,,		
JOURNAL	Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.		
COMMENT	Primer for synthesizing full-length cDNA and use thereof		
	Patent: JP 2002191363-A 13413 09-JUL-2002;		
	HELIX RESEARCH INSTITUTE		
	OS Homo sapiens (human)		
	PN JP 2002191363-A/13413		
	PD 09-JUL-2002		
	PF 28-JUL-2000 JP 2000280990		
	PI TOSHIO OTA,TAKAO ISOgai,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU		
	PI SAITO,		
	PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,		
	PI KEIICHI NAGAI,TETSUJI OTSUKI		
	PC		
	C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/		
	10,		
	PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00		
	CC Primer for synthesizing full-length cDNA and use thereof FH		
FEATURES			
source	FT CDS	Location/Qualifiers	Key
		(347)..(2131).	
ORIGIN			
	1..2131		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
Alignment Scores:			
Pred. No.:	2.07e-50	Length:	2131
Score:	501.00	Matches:	96
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-720-934-2_COPY_215_310 (1-96) x BD158570 (1-2131)			
Qy	1	TtpAlaValProGlnSerSerArgLeuLysTyArgGlnLeuPheAsnSerHisAspLys	20
Db	878	TGGGTGTTCTCCTCAGTCATCAAGACTGAAATACAGGCAATATTCATATGATCATGACAAA	937
Qy	21	ThrmSetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu	40

Dd	938	ACATAGTGTGACACTTAACAAGGTCGCCAACGAAGAACACTATTCTTAGCATAAGTTA	997
Oy	41	ProblinaIagInleuAlaseRIetRPaSnleuSerAsPIleaSpGlnaspGIYLeSieu	60
Dd	998	CCACAGGCTCAGCGTGCTCAATATGGAATCTTTCTCACATTGATCAAGATGGAAACTT	1057
Oy	61	ThralagluehelleleuAlamechIsleuileaspValAlametserGIYgInpro	80
Dd	1058	ACAGCAGAGGAATTTATCTCGGCATCGACCTCATTTATAGCTATGCTCGGCCAACCA	1117
Oy	81	leuprProvalleupProprogIUtyrllePProserPheaGagY	96
Dd	1118	CTGCACCTGCTCGCTCCAGAAATCAATTCACCTCTTTTAAAGA	1165
RESULT 8			
LOCUS	AKO27846		
DEFINITION	Homo sapiens cDNA FLJ14940 fis, clone PLACE1010942, highly similar to Homo sapiens intersectin long isoform (ITSN) mRNA.		
ACCESSION	AKO27846		
VERSION	AKO27846.1 GI:14042823		
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibaishi,T., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hirooka,S., Ishii,S., Kanai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Maehiro,Y. and Kanehorri,K. NEBO human cDNA sequencing project Unpublished 2 (bases 1 to 2131) Isogai,T. and Otsuki,T. Direct Submission Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomic@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.		
COMMENT			
FEATURES			
Source	Location/Qualifiers 1..2131 /organism="Homo sapiens" /mol_type="RNA" /db_xref="taxon:9606" /clone="PLACE1010942" /tissue_type="placenta" /clone_id="PLACE1" /note="cloning vector: pME18SFL3"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.07e-50	Length:	2131
Score:	501.00	Matches:	96
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-09-720-934-2_COPY_215_310 (1-96) x AKO27846 (1-2131)			
Oy	1	TPrAlaValProGlnSerSerArgLeuLeuLYTyArgGlnleuPheAnsSerHiASPLys	20
Dd	878	TGGGCTTCTTCAGTCATCATCAAGCTGAATATACGCGCAATTTATTCATATGTCAGACAA	937
Oy	21	ThrMetSerGIYHisleuThrglyProGlnAlaArgThrIleleuMetGlnSerSerIeu	40

Db 938 ACTATGATGACACTTACAGCTCCACAGCAAAATATCTTATGACGCAAGTTA 997
Qy 41 ProglinalaginLeuA1Ser11eTnpasnLeuSerAap11aaspG1aasp01yysleu 60
Db 998 CCACAGGCTCAGCTGCTTCAATATGCAATCTTTCTGACATGATCAAGATGCAAACTT 1057
Qy 61 Thr1agiug1uphe1leuA1aMeth1sleu1leaspVal1a1aMetSerG1yginPro 80
Db 1058 ACAGAGAGGAATTATATCTGCAATGACACCTCATGATGATGATGCTGGCCACCA 1117
Qy 81 LeuProP-oval1leuProProG1uTyr1lleProProSerPheArgArg 96
Db 1118 CTGCCACCTGCTCTCCCTCCAGAAATACATTCACCTTCTTTAGAGA 1165
RESULT 9
LOCUS AK074554 2131 bp mRNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens cDNA FLJ90073 fis, clone HEMBA1004110, highly similar
to Homo sapiens interseotin short form mRNA.
ACCESSION AK074554.1 GI:22760070
VERSION AK074554.1 GI:22760070
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1
1 Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2131)
AUTHORS Isogai,T. and Otsuki,T.
JOURNAL Direct Submission
TITLE Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
SOURCE location/Qualifiers
1..2131
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBA1004110"
/tissue_type="whole embryo, mainly head"
/clone_id="HEMBA1"
/dev_stage="embryo, 10 weeks"
/note="Cloning vector: pME18FLJ3"
ORIGIN
Alignment Scores: 2.07e-50 Length: 2131
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x AK074554 (1-2131)
Qy 1 TPpAlaValProGlnSerSerArgLeuYrYrArgGlnLeuPheAsnSerHisAspIys 20

Db 847 TGGGCTGTTCTCAGTCATCAAGACTGAATATACAGCAATATTCATATGATGACAA 906
Qy 21 ThnMetSerG1yH1sleuThrG1yProGlnA1aargThr1lleuMetGlnSerSer1eu 40
Db 907 ACTATGATGACACTTACAGCTCCACAGCAAAATATCTTATGACGCAAGTTA 966
Qy 41 ProglinalaginLeuA1Ser11eTnpasnLeuSerAap11aaspG1aasp01yysleu 60
Db 967 CCACAGGCTCAGCTGCTTCAATATGCAATCTTTCTGACATGATCAAGATGCAAACTT 1026
Qy 61 Thr1agiug1uphe1leuA1aMeth1sleu1leaspVal1a1aMetSerG1yginPro 80
Db 1027 ACAGAGAGGAATTATATCTGCAATGACACCTCATGATGATGATGCTGGCCACCA 1086
Qy 81 LeuProP-oval1leuProProG1uTyr1lleProProSerPheArgArg 96
Db 1087 CTGCCACCTGCTCTCCCTCCAGAAATACATTCACCTTCTTTAGAGA 1134
RESULT 10
LOCUS BC058925 2199 bp mRNA linear PRI 05-NOV-2003
DEFINITION Homo sapiens interseotin 1 (SH3 domain protein), mRNA (cDNA clone
IMAGE:4443129), partial cds.
ACCESSION BC058925
VERSION BC058925.1 GI:37589134
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2199)
1 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Aleschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schaefer,T.E., Brownstein,M.J., Udell,T.B., Toshyiki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullenb,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,J.J., Hallyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J.J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 2199)
AUTHORS Strausberg,R.
JOURNAL Direct Submission
TITLE Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapsb-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Genome Sequence Center,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Nataja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
Series: IRAK Plate: 119 Row: B Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796.

FEATURES

source

1. 2199

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4443129"

/tissue_type="Liver, adenocarcinoma"

/clone_lib="NIH MGC_90"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

gene

1. .>2199

/gene="ITSN1"

/note="synonyms: ITSN, SH3D1A, SH3P17"

/db_xref="LocusID:6453"

/db_xref="MIM:602442"

CDS

296. .>2199

/gene="ITSN1"

/codon_start=1

/product="ITSN1 protein"

/protein_id="AAH58925.1"

/db_xref="GI:37589135"

/db_xref="LocusID:6453"

/db_xref="MIM:602442"

misc_feature

335. .619

/gene="ITSN1"

/note="EH; Region: Eps15 homology domain"

/db_xref="CDT:smart00027"

misc_feature

935. .1171

/gene="ITSN1"

/note="EH; Region: Eps15 homology domain"

/db_xref="CDT:smart00027"

ORIGIN

Alignment Scores:

Pred. No.: 2.15e-50 Length: 2199
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x BC058925 (1-2199)

QY 1 TTPALVALPGLINSERATGLeuLYTFAAGLIneuPheasnsERH:asPLYs 20

Db 938 TGGGCTGTTCTCAGTCATCAAGCTCAATATTCATGCTCATGACAA 997

QY 21 ThmetSerGIYHISLeuThrgLYProGInalAargThrlleuWetGInserSerLeu 40

Db 998 ACTATGAGTGACATTAAACAGGTCCCAAGCAAGAACTATCTTATGACGTCAAGTTA 1057

QY 41 ProGInalAclneualaseRllETTPaInleuserasPtleaspGInasgLYsleu 60
Db 1058 CCACAGGCTCAGCTGCCTCCATATGGAATCTTTGACATTCATCAAGATGCAAACTT 1117

QY 61 ThrAlaGluGluPheleuAlamethISleuileasvAlaImetSerGIYgInPro 80
Db 1118 ACAGCAGAGGAATTTATCTCGCAATGACCTCATATTAGTACTATGCTCGGCCAACCA 1177

QY 81 LeuProPreValleuProProGIuYrileProPserPheargatg 96
Db 1178 CTGCACCTGCTCCTGCCTCCAGATACCTTCCTTTTAGAAGA 1225

RESULT 11

AF132478

LOCUS AF132478 3723 bp mRNA linear ROD 09-MAR-1999

DEFINITION Mus musculus Eps1 protein mRNA, complete cds.

ACCESSION AF132478

VERSION AF132478.1 GI:4378884

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PubMed

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

1. 3723

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

1. 3642

/codon_start=1

/product="Eps1 protein"

/protein_id="FAD19746.1"

/db_xref="GI:4378885"

/translation="MAQFPPTPGSGSLDVAITVEERAKHDQOFLSLKPIAGTITGDA
RNFFPQSGLPQVLAQIMALADNNGRMDQVFFSIAMKLIKIKLQGYOLPSTLPVM
KOQPAVSISSAPAFGIGIASMPPLTVAVPVGMSITVVGSPPLVSSVPPAAVPLAN
GAPVITQPLPAPAHAPATWPKSSFSRSGSGSLNKLQAGSFDAASAPAEWAVP
QSRILKYROLFNSHDKTMSGHLTGPOAKITLIMOSSLPQOLASINMUSIDIDQGLTA
EEFTIAMHLIDVAMSGQPLPVPPIPPSPFRVSGSGSVISSTVDQRLPEBPV
LEDEOQLEKRLPVTPEDEKKNRBERGNLEKRRQALLQOQKEQERLAQERAEQ
RKREKROEOERKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLE
RNRROELLNORNKOEEDIVYLKAKKTLLEFELALNDKKHOLEKLODTRCLTTRRO
ELETSTKSRLEARIATITHLQOQLOESQOMIGRIIPKQILINDLKQYQONSIRDSIV
TIKRALEAKELARQHRLDQLEVEKETRSLQEDIDIFNNQLEKRLIHNKQLOKQKS
MEAEKRLKKK"

ORIGIN

Alignment Scores:

Pred. No.: 3.8e-50 Length: 3723

Score: 501.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 10
Matches: 96
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AF132478 (1-3723)

QY 1 TrpAlaValAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 643 TGGGCTGGCTCCTCAGTCAAGGCTGAATAACAGCGATTATTCACAGCACACACAAA 702
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 703 ACTATGATGACACTTACAGGCTCCCGAGCAAGACTATTTCTATCATCAATCAAGTTTA 762
QY 41 ProGlnAlaGlnLeuAlaSerIleThrAsnLeuSerAspIleAspGlnLysLeu 60
Db 763 CCCGAGGCTCAGCTGCTTCAATGAAATCTTCTGCACATGATCAGATGAAACATC 822
QY 61 ThrAlaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 823 ACTGACAGAGAAATTTATCTAGTCACTTAATGACCTTAATGATGATGCTGCTGACGCA 882
QY 81 LeuProProValLeuProGluTyrIleProProSerPheArgArg 96
Db 883 CTGCCGCGCTCTCTCCTCCAGAAATACATCCCTCTCTCTTCCAGAAAGA 930

RESULT 12
AF132672 3812 bp mRNA linear ROD 28-JUN-1999
LOCUS AF132672
DEFINITION Rattus norvegicus EH-domain/SH3-domain containing protein mRNA,
complete cds.

ACCESSION AF132672.1 GI:4838525
VERSION AF132672.1
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 3812)
Okamoto, M., Schoch, S., and Sudhof, T.C.
EHSH1/intersectin, a protein that contains EH and SH3 domains and
binds to dynamin and SNAP-25. A protein connection between
exocytosis and endocytosis?
J. Biol. Chem. 274 (26), 18446-18454 (1999)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 3812)
Okamoto, M., Schoch, S., and Sudhof, T.C.
Direct Submission
Submitted (02-MAR-1999) Center for Basic Neuroscience and HHMI,
Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX
75235, USA

FEATURES
source
1. 3812
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
41..3481
/note="EHSH1"
/codon_start=1
/product="EH-domain/SH3-domain containing protein"
/protein_id="AF132672.1"
/db_xref="gi:4838525"

CDS
1 (bases 1 to 4025)
Okamoto, M., Schoch, S., and Sudhof, T.C.
EHSH1/intersectin, a protein that contains EH and SH3 domains and
binds to dynamin and SNAP-25. A protein connection between
exocytosis and endocytosis?
J. Biol. Chem. 274 (26), 18446-18454 (1999)
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 4025)
Okamoto, M., Schoch, S., and Sudhof, T.C.
Direct Submission
Submitted (11-FEB-1999) Center for Basic Neuroscience and HHMI,
Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX
75235, USA

ORIGIN

Alignment Scores:

Pred. No.: 3.9e-50 Length: 3812
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AF132672 (1-3812)

QY 1 TrpAlaValAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 683 TGGGCTGGCTCCTCAGTCAAGGCTGAATAACAGCGATTATTCACAGTACACACAG 742
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 743 ACCATGATGACACTTAAACAGGCTCCCGAGCAAGACTATTTCTATCATCAATCAAGTTTA 802
QY 41 ProGlnAlaGlnLeuAlaSerIleThrAsnLeuSerAspIleAspGlnLysLeu 60
Db 803 CCCGAGGCTCAGCTGCTTCAATGAAATCTTCTGCACATGATCAGATGAAAGCTC 862
QY 61 ThrAlaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 863 ACCGACAGAAATTTATCTAGTCACTTAATGACCTTAATGATGATGCTGCTGACGCA 922
QY 81 LeuProProValLeuProGluTyrIleProProSerPheArgArg 96
Db 923 CTGCCGCGCTCTCTCCTCCAGAAATACATCCCTCTCTCTTACAGAA 970

RESULT 13

LOCUS AF127798 4025 bp mRNA linear ROD 28-JUN-1999
DEFINITION Rattus norvegicus EH- and SH3-domain containing protein EHSH1 mRNA,
complete cds.

ACCESSION AF127798
VERSION AF127798.1 GI:4835852
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 4025)
Okamoto, M., Schoch, S., and Sudhof, T.C.
EHSH1/intersectin, a protein that contains EH and SH3 domains and
binds to dynamin and SNAP-25. A protein connection between
exocytosis and endocytosis?
J. Biol. Chem. 274 (26), 18446-18454 (1999)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 4025)
Okamoto, M., Schoch, S., and Sudhof, T.C.
Direct Submission
Submitted (11-FEB-1999) Center for Basic Neuroscience and HHMI,
Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX
75235, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 4025)
Okamoto, M., Schoch, S., and Sudhof, T.C.
Direct Submission
Submitted (11-FEB-1999) Center for Basic Neuroscience and HHMI,
Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX
75235, USA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 11:41:07 ; Search time 296.319 Seconds
(without alignments)
1376.313 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 WAWPQSSRLKXRLNLSHDK.....SQQLPVPVLPPEYIPPSFR 96

Scoring table:
BL0SUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O/cnrm2.1/USPTO.spool_p/US09700934/runat.30062004.064539.13422/app.query.fasta_1.1386
-DB=N Geneseg 297and4 -QPM=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPT=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09700934 @CGM 1.1 1596 @runat.30062004.064539.13422 -NCPU=6 -ICPU=3
-NO MMAP -LARGEJITTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseg 297and4:*

1: genesegq1980s:.*
2: genesegq1990s:.*
3: genesegq2000s:.*
4: genesegq2001as:.*
5: genesegq2001bs:.*
6: genesegq2002s:.*
7: genesegq2003as:.*
8: genesegq2003bs:.*
9: genesegq2003cs:.*
10: genesegq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	501	100.0	1676	4	AAK94611 Human CDN
2	501	100.0	2079	2	AAZ34573 Human SH3
3	501	100.0	2131	4	AAK94133 Human ful
4	501	100.0	2131	4	AAH16578 Human CDN
5	501	100.0	3723	3	AAZ39009 Mouse Ese
6	501	100.0	5082	3	AAZ39008 Mouse Ese
7	501	100.0	5144	3	AAZ39025 Mouse Ese
8	501	100.0	5195	2	AAZ34572 Human SH3

9	501	100.0	5199	2	AAZ34570 Human SH3
10	501	100.0	5458	2	AAZ34571 Human SH3
11	501	100.0	5738	3	AAZ39024 Mouse Ese
12	496	99.0	548	4	AA180000 Human pol
13	489	97.6	3319	4	ABK43498 DNA encod
14	489	97.6	3466	4	AA163825 Human pol
15	416	83.0	7435	5	AA584763 DNA encod
16	349	69.7	2017	3	AAAG9762 Human ova
17	349	69.7	2017	6	ABN72656 Ovarian c
18	349	69.7	2017	8	ADA08821 Human ova
19	349	69.7	3593	3	AAZ39011 Mouse Ese
20	349	69.7	4625	3	AAZ39010 Mouse Ese
21	349	69.7	4975	3	AAZ39027 Mouse Ese
22	349	69.7	6014	3	AAZ39026 Mouse Ese
23	345	68.9	4447	4	AA502055 DNA encod
24	345	68.9	5828	6	AA147247 Allergic
25	345	68.9	6103	4	AAK52332 Human pol
26	337	67.3	346	3	AA444038 Human sec
27	334	66.7	955	4	AA163826 Human pol
28	334	66.7	955	4	AA531620 DNA encod
29	334	66.7	955	4	ABK43501 DNA encod
30	292	58.3	3981	4	AB101995 Drosophila
31	292	58.3	7225	4	AB101994 Drosophila
32	280	55.9	831	4	AAH03435 Human CDN
33	280	55.9	1329	4	AAH15260 Human CDN
34	208	41.5	480	3	AAAG9750 Human ova
35	208	41.5	480	6	ABN72644 Ovarian c
36	208	41.5	480	8	ADA08809 Human ova
37	208	41.5	531	3	AAAG9695 Human ova
38	208	41.5	531	6	ABN72589 Ovarian c
39	208	41.5	531	8	ADA08754 Human ova
40	204.5	40.8	3033	2	AAV13999 DNA encod
41	204.5	40.8	3033	2	AAV13999 Murine ep
42	204.5	40.8	3033	2	AAH04192 Murine ep
43	203.5	40.6	2666	4	AAH18433 Human CDN
44	203.5	40.6	2774	3	AAZ96752 Nuclear t
45	203.5	40.6	2782	4	AAH19248 Human sec

ALIGNMENTS

RESULT 1
AAK94611 standard; cDNA; 1676 BP.
ID AAK94611
AC AAK94611;
XX
XX
DT 07-NOV-2001 (first entry)
XX
XX
DE Human cDNA, SEQ ID NO: 3565.
XX
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN EP1130094-A2.
XX
XX
PD 05-SEP-2001.
XX
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX
PA (HELI-) HELIX RES INST.
XX
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX
WP1; 2001-524255/58.
XX
XX
P-PSDB; AAM93676.
XX
XX

PT 830 Primers useful for synthesizing full length cDNA clones and their use
in genetic manipulation.
XX
XX
PS Disclosure; SEQ ID NO 3565; 1380bp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
clones. 830 cDNA molecules encoding a human protein have been isolated
and nucleotide sequences of 5' and 3' ends of the cDNA molecules have
been determined. Primers for synthesizing the full length cDNA are useful
for clarifying the function of the protein encoded by the cDNA. The full
length clones were obtained by construction of full length enriched cDNA
libraries that were synthesized by the oligo-capping method. The primers
enable the production of the full length cDNA easily without any special
methods. The present sequence is a human cDNA provided in the
specification. Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in CD-ROM format directly
from EPO
XX
SQ Sequence 1676 BP; 485 A; 381 C; 458 G; 352 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,028-52 Length: 1676
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x AA394611 (1-1676)
QY 1 TTPALAVAlProGInSerSerArgLeuLysTYRArgGInLeuPheAsnSerHisAspLys 20
Db 906 TGGGCTGTTCCTCACTCATCAAGACTGAATATCAGGCAATATTATTAATAGTCATACAA 965
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 966 ACTATGAGTGAACACTTAAACAGTCTCCCAAGCAAGAACTATCTTATCAGTCAAGTTTA 1025
QY 41 ProGlnAlaGlnLeuAlaSerIleThrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 1026 CCACAGGCTTCAGCTGCTTCATATGGAATCTTCTGACATTGATCAAGATGGAATCTT 1085
QY 61 ThrAlaGlnIleupheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 1086 ACAGAGAGGAAATTATCTCGCAATGACCTCATTTGATGATGATCTCGGCCAACCA 1145
QY 81 LeuProProValLeuProProGlnIleuTyrlleProProSerPheArgArg 96
Db 1146 CTGCCACCTGTCCTGCCCTCCAGAAATACATTCACCTCTTTTAGAGAGA 1193
RESULT 2
AA234573
ID AA234573 standard; cDNA; 2079 BP.
XX
AC AA234573;
XX
DT 01-FEB-2000 (first entry)
XX
XX Human SH3D1A cDNA clone 5.
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 136..2079
FT CDS
FT /tag= a
XX
XX PN W09953062-A2.

XX
PD 21-OCT-1999.
XX
XX
PF 16-APR-1999; 99WO-US008371.
XX
PR 16-APR-1999; 98US-0082007P.
XX
XX (CEDA-) CEPARS SINAI HEALTH SYSTEM.
XX
XX
PI Korenberg JR, Chen X;
XX
XX MPI: 1999-633829/54.
DR P-PsDB; AA32157.
XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
XX
PS Claim 2; Fig 12; 99pp; English.

CC This is the nucleotide sequence of a cDNA clone, termed clone 5,
CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
CC contributes to the development of platelets and the pathogenesis of
CC leukaemias, both in general and in particular those involving the
CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
CC region for low platelets on chromosome 21. Sequencing of 5 different
CC sizes of cDNA clone from fetal brain (see AA234570-74) suggests that at
CC least 3 isoforms exist. The invention provides methods for the diagnosis
CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
CC association of gains in chromosome 21 with leukaemias, neural
CC abnormalities, dysfunctions and disorders including brain malformations
CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
CC and colpocephaly. Methods are also provided for: suppressing cells unable
CC to regulate themselves; screening for a somatic alteration in the SH3D1A
CC gene; monitoring the progress and adequacy of a treatment; monitoring
CC tumour risk progress or megakaryocytic abnormality; myeloproliferative
CC disorder; haematopoietic disorder; platelet disorder or leukaemia; and
CC treatment of a subject (including a prenatal subject) having
CC megakaryocytic abnormality, myeloproliferative disorder, platelet
CC disorder, leukaemia or neural disorder using a nucleic acid that
CC expresses SH3D1A or its antisense nucleic acid
XX

SO Sequence 2079 BP; 700 A; 451 C; 495 G; 433 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,378-52 Length: 2079
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AA234573 (1-2079)

QY 1 TTPALAVAlProGInSerSerArgLeuLysTYRArgGInLeuPheAsnSerHisAspLys 20
Db 778 TGGGCTGTTCCTCACTCATCAAGACTGAATATCAGGCAATATTATTAATAGTCATACAA 837
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 838 ACTATGAGTGAACACTTAAACAGTCTCCCAAGCAAGAACTATCTTATCAGTCAAGTTTA 897
QY 41 ProGlnAlaGlnLeuAlaSerIleThrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 898 CCACAGGCTCAGCTGCTTCATATGGAATCTTCTGACATTGATGATCAAGATGGAATCTT 957
QY 61 ThrAlaGlnIleupheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 958 ACAGCAGAGAAATTATCTCGCAATGACCTCATTTATGATGATGATCTCGGCCAACCA 1017
QY 81 LeuProProValLeuProProGlnIleuTyrlleProProSerPheArgArg 96

Db 1018 CTGGCACCTGTCCTCGCTCCAGAAATACATTCACCTTCTTTAGAGA 1065

RESULT 3

AAK94139
ID AAK94139 standard; cDNA; 2131 BP.
XX
AC AAK94139;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 2646.
XX
KM Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR P-PSDB; AAM93229.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 2646; 1380bp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a full length human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in CD-ROM format directly
CC from EPO
XX
SQ Sequence 2131 BP; 694 A; 469 C; 526 G; 442 T; 0 U; 0 Other;
XX

Alignment Scores:

Pred. No.:	1,41e-52	Length:	2131
Score:	501.00	Matches:	96
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_215_310 (1-96) x AAK94139 (1-2131)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 847 TGGGCTGTCCTCCTCATCAAGACTGAATATACAGGCAATTATTCAATAGCATGACAA 906
QY 21 ThwTserGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 907 ACTATGAGTGACACTTAACAGGATCCCAAGCAAGAACTATCTTATGACAGTCAAGTTTA 966
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60

Db 967 CCAAGGCTCAGCTGGCTTAAATATGAAATCTTTCTGACTGATCAAGATGAAAACCTT 1026

QY 61 ThwAlaGluIubPheIleuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 1027 ACAGCAGAGGAATTTATCCGCGCAATGACACTCATTTAGTATGTCGCGCAACCA 1086

QY 81 LeuProProValLeuProProGlyTyrIleProProSerPheArgArg 96
Db 1087 CTGGCACCTGTCCTCGCTCCAGAAATACATTCACCTTCTTTAGAGA 1134

RESULT 4

AAH16578
ID AAH16578 standard; cDNA; 2131 BP.
XX
AC AAH16578;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:15658.
XX
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 26-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Makamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 15658; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dr primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH16628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

```

XX SQ Sequence 2131 BP; 691 A; 448 C; 572 G; 420 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.41e-52 Length: 2131
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x AAH16578 (1-2131)
QY 1 TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 878 TGGGCTGTTCCTCACTGATCAAGACTGAATACAGGCAATTATTAATATGATATACAAA 937
QY 21 ThrmSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
DB 938 ACTATGATGAGCACTTAACAGGTCCCAAGCAAGAACATTTCTATGAGTCAAGTTTA 997
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysIeu 60
DB 998 CCACAGGCTCAGCTGCTTCATATGAAATCTTTCTGACATTGATCAAGATGCAAAACTT 1057
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1058 ACAGAGAGAGAAATTATCTCTGCAATGCACTCATTTGATGATGCTGAGCCACCA 1117
QY 81 LeuProValIleuProGluTyrIleProPheArgArg 96
DB 1118 CTGCCACCTGTCTCTCCATATACATTCACCTTCTTTAGAAGA 1165
RESULT 5
AAZ39009
ID AAZ39009 standard; cDNA, 3723 BP.
XX AC AAZ39009;
XX DT 28-FEB-2000 (first entry)
XX DE Mouse Esei coding sequence.
XX KW Mouse; murine; Esei; Ees2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX OS Mus sp.
XX PN WO9955728-A2.
XX PD 04-NOV-1999.
XX PF 27-APR-1999; 99WO-CA000375.
XX PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX PA (HSCR-) HSC RBS & DEV LP.
XX PI Egan SE, Wang W, Sengar A;
XX WP1: 2000-052802/04.
DR P-PSDB; AAY57444.
XX PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of
XX endocytosis, used e.g. for treating cancer or preventing viral infection.
XX PS Claim 6; Page 40-42; 99pp; English.
XX CC The present sequence encodes mouse Esei. The present invention
CC specifically describes mammalian Esei and 2 proteins (I) and their splice

```

```

CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Ees1s protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Esei is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Ees1s complex, then binding dynamn to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signaling, tissue development or synaptic transmission
XX SQ Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2.97e-52 Length: 3723
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x AAZ39009 (1-3723)
QY 1 TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 643 TGGGCTGTTCCTCACTGATCAAGGCTGAATACAGGCAATTATTAACAGCCACACAAA 702
QY 21 ThrmSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
DB 703 ACTATGATGAGCACTTAACAGGTCCCAAGCAAGAACATTTCTATGATCAAGTTTA 762
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysIeu 60
DB 763 CCCACAGGCTCAGCTGCTTCATATGAAATCTTTCTGACATTGATCAAGATGCAAAACTT 822
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 823 ACTGAGAGAGAAATTATCTCTGCAATGCACTCATTTGATGATGCTGAGCCACCA 882
QY 81 LeuProValIleuProGluTyrIleProPheArgArg 96
DB 883 CTGCCGCCGCTGTCTCCATATACATTCCTCTTCTTCAAGAAGA 930
RESULT 6
AAZ39008
ID AAZ39008 standard; cDNA, 5082 BP.
XX AC AAZ39008;
XX DT 28-FEB-2000 (first entry)
XX DE Mouse Esei full length cDNA sequence.
XX KW Mouse; murine; Esei; Ees2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX OS Mus sp.
XX PN WO9955728-A2.
XX PD 04-NOV-1999.

```

```
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
XX (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
DR WPI; 2000-052802/04.
XX P-PSDB; AAY57444.
XX
PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
XX endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS Claim 6; Page 38-40; 99pp; English.
XX
XX The present sequence encodes mouse Ese1. The present invention
XX specifically describes mammalian Ese1 and 2 proteins (1) and their splice
XX variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)
XX are involved in regulation of clathrin-mediated endocytosis (as a complex
XX with Esp15 protein), vesicular trafficking and actin cytoskeleton.
XX Generally (1) (or its (ant)agonists, mimetics, fragments and inactive
XX mutants); (1)-specific antibodies (Ab); sequences antisense to the (1)
XX polynucleotide; agents that downregulate expression of Ese genes or
XX antagonists of an Ese binding partner are used to treat diseases
XX associated with undesirable endocytosis and resulting changes in cellular
XX function. Particularly overexpression of Ese1 is used to block clathrin-
XX mediated endocytosis in vivo or in cell cultures, while administration of
XX (1) is used to promote endocytosis of selected cells. (Ant)agonists of
XX (1) or Ab are used to suppress abnormal proliferation of cells that can
XX be stimulated to proliferate by a growth factor receptor; and similar
XX compounds (also inactive Ese mutants) can be used to prevent viral
XX infection. Endocytosis may also be regulated, in vivo or in cell
XX cultures, by forming an Ese-Esp15 complex, then binding dynamin to the
XX complex. Generally conditions that can be treated include cancer;
XX abnormal cell division or migration; viral infection; or abnormal
XX receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.5e-52 Length: 5082
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 3
US-09-720-934-2_COPY_215_310 (1-96) x AAZ39008 (1-5082)
QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAenSerHisAspLys 20
Db TGGGCTGTGCTCAGTCATCAAGCTGAATATACAGGAGATTATTCAACAGCAGCAAA 960
QY 21 ThMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerLeu 40
Db ACTATGTGTGACACTTAACAGGCTCCAGGACAGAACTATTCACAGCAATCAAGTTA 1020
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db CCCAGGCTCAGCTGGCTTCAATATGGAATCTTTTCACATTGATCAAGATGAAAACTC 1080
QY 61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db ACTGCAGAAAGAAATTATCTAGCTATGCACTTAATGATGTGCGATGCTGAGCAGCA 1140
QY 81 LeuProProValLeuProProGlnTyrIleProProSerPheArgang 96
Db 1141 CTGGCGCCGCTGCTGCTCCAGAAATACATCCCTCTCTTCAGAAAGA 1188
RESULT 7
AAZ39025
```

```
ID AAZ39025 standard; cDNA; 5144 BP.
XX
XX AAZ39025;
AC
XX 28-FEB-2000 (first entry)
XX
XX Mouse Ese1 coding sequence.
XX
XX Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
XX regulation; actin cytoskeleton; detection; cancer; infection;
XX EH-domain and SH3-domain regulator of endocytosis; anticancer;
XX antiproliferative; antiviral; ss.
XX
XX Mus sp.
XX
XX MO9955728-A2.
XX
XX 04-NOV-1999.
XX
XX 27-APR-1999; 99WO-CA000375.
XX
XX 27-APR-1998; 98CA-02230201.
XX 05-FEB-1999; 99US-0118739P.
XX
XX (HSCR-) HSC RES & DEV LP.
XX
XX Egan SE, Wang W, Sengar A;
XX
XX WPI; 2000-052802/04.
XX P-PSDB; AAY57449.
XX
XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
XX endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
XX Claim 6; Page 59-62; 99pp; English.
XX
XX The present invention specifically describes mammalian Ese1 and 2
XX proteins (1) and their splice variants (Ese = EH-domain and SH3-domain
XX regulator of endocytosis). (1) are involved in regulation of clathrin-
XX mediated endocytosis (as a complex with Esp15 protein), vesicular
XX trafficking and actin cytoskeleton. Generally (1) (or its (ant)agonists,
XX mimetics, fragments and inactive mutants); (1)-specific antibodies (Ab);
XX sequences antisense to the (1) polynucleotide; agents that downregulate
XX expression of Ese genes or antagonists of an Ese binding partner are used
XX to treat diseases associated with undesirable endocytosis and resulting
XX changes in cellular function. Particularly overexpression of Ese1 is used
XX to block clathrin-mediated endocytosis in vivo or in cell cultures, while
XX administration of (1) is used to promote endocytosis of selected cells.
XX (Ant)agonists of (1) or Ab are used to suppress abnormal proliferation of
XX cells that can be stimulated to proliferate by a growth factor receptor;
XX and similar compounds (also inactive Ese mutants) can be used to prevent
XX viral infection. Endocytosis may also be regulated, in vivo or in cell
XX cultures, by forming an Ese-Esp15 complex, then binding dynamin to the
XX complex. Generally conditions that can be treated include cancer;
XX abnormal cell division or migration; viral infection; or abnormal
XX receptor signalling, tissue development or synaptic transmission. The
XX present sequence represents mouse Ese1 coding sequence
XX
SQ Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.58e-52 Length: 5144
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 3
US-09-720-934-2_COPY_215_310 (1-96) x AAZ39025 (1-5144)
QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAenSerHisAspLys 20
Db TGGGCTGTGCTCAGTCATCAAGCTGAATATACAGGAGATTATTCAACAGCAGCAAA 702
```

QY 21 ThrMetSerGIYHISLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 DB 703 ACTATGAGTGGACACTTAAAGTCCCAAGCAAACTATTCTCATCAATCAAGTTTA 762
 QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyIysIeu 60
 DB 763 CCCGAGGCTGAGCTGGCTTCATATGAGATCTTTTGACATTGATCAAGATGGAAATCTC 822
 QY 61 ThrAlaGluGluPheIleLeuAlaMetHisIleuIleAspValAlaMetSerGlyGlnPro 80
 DB 823 ACTGAGAGAAATTTATCTAGCTATGACACCTAATTGATGTTCATGTCTGGTCAGCA 882
 QY 81 LeuProPValIleuProGlnIuTyrIleProPserPheArg 96
 DB 883 CTGCGCCGCTGCTGCTCCAGAAATACCTCTCTCTTCAAGAGA 930
 RESULT 8
 ID AA234572 standard; cDNA; 5195 BP.
 AC AA234572;
 XX
 XX 01-FEB-2000 (first entry)
 DT
 XX Human SH3D1A cDNA clone 11.
 DE
 XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 239..3886
 FT /*tag= a
 FT
 XX W09953062-A2.
 PN
 XX 21-OCT-1999.
 PD
 XX 16-APR-1999; 99W0-US008371.
 PF
 XX 16-APR-1998; 98US-0082007P.
 PR
 XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 PA
 XX Korenberg JR, Chen X;
 PI
 DR WPI, 1999-633829/54.
 DR P-PSDB; AAY32156.
 DR
 XX Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 PS Claim 2; Fig 10; 99pp; English.
 PS
 XX This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukaemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukaemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,

CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4,64e-52 Length: 5195
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 2
 US-09-720-934-2_COPY_215_310 (1-96) x AA234572 (1-5195)
 QY 1 TrrAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspIys 20
 DB 881 TGGGCTGTTCTCTCAGTCATCAAGACTGAATACAGGCAATTAATCAATAGTCATGACAAA 940
 QY 21 ThrMetSerGIYHISLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 DB 941 ACTATGAGTGGACACTTAAAGTCCCAAGCAAACTATTCTTATCAAGTCAAGTTTA 1000
 QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyIysIeu 60
 DB 1001 CCACAGGCTCAGCTGGCTTCATATGAGATCTTTCTGACATTGATCAAGATGGAAACTT 1060
 QY 61 ThrAlaGluGluPheIleLeuAlaMetHisIleuIleAspValAlaMetSerGlyGlnPro 80
 DB 1061 ACAGCAGAGAAATTTATCTGCAATGACCTCTTGTATGTAGTCATGTGGCCACACA 1120
 QY 81 LeuProPValIleuProGlnIuTyrIleProPserPheArg 96
 DB 1121 CTGCCACTGTGCTGCTCCAGAAATACCTCTCTTTTAGAGA 1168
 RESULT 9
 ID AA234570 standard; cDNA; 5199 BP.
 AC AA234570;
 XX
 XX 01-FEB-2000 (first entry)
 DT
 XX Human SH3D1A cDNA.
 DE
 XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 208..3642
 FT /*tag= a
 FT
 XX W09953062-A2.
 PN
 XX 21-OCT-1999.
 PD
 XX 16-APR-1999; 99W0-US008371.
 PF
 XX 16-APR-1998; 98US-0082007P.
 PR
 XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 PA

XX Korenberg JR, Chen X;
 XX WPI; 1999-633829/54.
 DR P-PSDB; AAY32154.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 5; 99pp; English.

CC This is the nucleotide sequence of full-length cDNA corresponding to a
 CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the
 CC development of platelets and the pathogenesis of leukemias, both in
 CC general and in particular those involving the megakaryocytic lineage. The
 CC SH3D1A gene maps to the small candidate region for low platelets on
 CC chromosome 21. Sequencing of 5 different sizes of cDNA clone (see
 CC AA234570-74) suggests that at least 3 isoforms exist. The invention
 CC provides methods for the diagnosis and treatment of megakaryocytic
 CC abnormality, myeloproliferative disorder, platelet disorder, acute
 CC leukemia, neural disorders, thrombocytopenia, platelet disorder on
 CC chromosome 21, low platelets in deletion for 21, association of gains in
 CC chromosome 21 with leukemias, neural abnormalities, dysfunctions and
 CC disorders including brain malformations and corresponding cognitive
 CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
 CC also provided for: suppressing cells unable to regulate themselves;
 CC screening for a somatic alteration in the SH3D1A gene; monitoring the
 CC progress and adequacy of a treatment; monitoring tumour risk progress or
 CC megakaryocytic abnormality, myeloproliferative disorder, haematopoietic
 CC disorder, platelet disorder or leukemia; and treatment of a subject
 CC (including a prenatal subject) having megakaryocytic abnormality,
 CC myeloproliferative disorder, platelet disorder, leukemia or neural
 CC disorder using a nucleic acid that expresses SH3D1A or its antisense
 CC nucleic acid
 XX

XX Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.64e-52 Length: 5199
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AA234570 (1-5199)

QY 1 TTPAlaValProGlnSerSerArgLeuYsTYrArgGlnLeuPheAsnSerHisAspLys 20
 DB 850 TGGGCTGTCTCCTCAGTCATCAAGACGAAATACAGGCAATATTATCAATGTCATGCAAA 909
 QY 21 ThMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
 DB 910 ACTATAGTGGACACTTAACAGGTCCGCCAAGCAAGAACTATCTTATGACGTCAAGCTTAA 969
 QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGlnAspGlyLysLeu 60
 DB 970 CCAAGGCTCAAGCTGCTTCAATATGAATCTTTCGACCTTGATCAAGATGGAAAACTT 1029
 QY 61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 DB 1030 ACAGCAGAGAAATTATTCCTGGCAATGCACTCATTTGATGATGCTATGCTGGCCAAACCA 1089
 QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
 DB 1090 CTGCCACCTGTCTGCTCCAGAAATACATTCACCTCTTTTGAAGA 1137

RESULT 10
 AA234571

XX AA234571 standard; cDNA; 5458 BP.

XX
 AC
 XX AA234571;

DT 01-FEB-2000 (first entry)
 XX
 XX Human SH3D1A cDNA clone 21.
 DE
 XX SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 267..3929
 FT /*tag= a

XX WO9553062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WO-US008371.
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX

PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.

PI Korenberg JR, Chen X;
 XX
 DR WPI; 1999-633829/54.
 DR P-PSDB; AAY32155.

PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 8; 99pp; English.

CC This is the nucleotide sequence of full-length cDNA (clone 21)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukemia, neural disorders, thrombocytopenia,
 CC association of gains in chromosome 21, low platelets in deletion for 21,
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX

SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.95e-52 Length: 5458
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AA234571 (1-5458)

QY 1 TTPAlaValProGlnSerSerArgLeuYsTYrArgGlnLeuPheAsnSerHisAspLys 20

DB 909 TGGGCTGTCTCTAGTATCAAGACGTAATACAGGCAATTATTCATATGATCAACAAA 968
QY 21 ThwterSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerIeu 40
DB 969 ACTATGAGTGACACTTAACAGGCTCCCAAGCAAGAACTATTTATGACATCAAGTTTA 1028
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLeu 60
DB 1029 CCACAGGCTCAGCTGGCTTCATATGAAATCTTTCTGACATTGATCAAGATGAAAACTT 1088
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1089 ACACGAGAGAAATTTATCTGATGATGACCTCATTTGATGCTATGTCTGGCCAAACA 1148
QY 81 LeuProValLeuProGlnTyrIleProSerPheArgArg 96
DB 1149 CTGCCACCTGCTGCTCCAGAAATACATTCACCTCTTTAGAGA 1196
RESULT 11
AAZ39024
ID AAZ39024 standard; cDNA; 5738 BP.
AC AAZ39024;
XX 28-FEB-2000 (first entry)
XX Mouse Esei1. cDNA sequence.
DE Mouse; murine; Esei1; Esei2; endocytosis; vesicular trafficking;
KM regulation; actin cytoskeleton; detection; cancer; infection;
KM EH-domain and SH3-domain regulator of endocytosis; anticancer;
KM antiproliferative; antiviral; ss.
OS Mus sp.
XX WO9955728-A2.
PN 04-NOV-1999.
XX 27-APR-1999; 99WO-CA000375.
PF 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX (HSCR-) HSC RES & DEV LP.
PA Egan SE, Wang W, Sengar A;
XX WPI: 2000-052802/04.
DR P-PSDB; AA57449.
XX New nucleic acid encoding Esei and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX Claim 6; Page 56-59; 99pp; English.
XX The present invention specifically describes mammalian Esei and 2
CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
CC regulator of endocytosis). (I) are involved in regulation of clathrin-
CC mediated endocytosis (as a complex with Esp15 protein), vesicular
CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
CC mimetics, fragments and inactive mutants); (I)-specific antibodies,
CC sequences antisense to the (I) polynucleotide; agents that downregulate
CC expression of Ese genes or antagonists of an Ese binding partner are used
CC to treat diseases associated with undesirable endocytosis and resulting
CC changes in cellular function. Particularly overexpression of Esei is used
CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
CC administration of (I) is used to promote endocytosis of selected cells.
CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
CC cells that can be stimulated to proliferate by a growth factor receptor;
CC and similar compounds (also inactive Ese mutants) can be used to prevent
CC viral infection. Endocytosis may also be regulated, in vivo or in cell

CC cultures, by forming an Ese-Esp15 complex, then binding dynamin to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission. The
CC present sequence represents mouse Esei1 cDNA sequence
XX
SQ Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.3e-52 Length: 5738
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-720-934-2_copy_215_310 (1-96) x AAZ39024 (1-5738)
QY 1 TrpAlaValProGlnSerSerArgLeuIleTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 902 TGGGCTGTCTCTAGTATCAAGGCTGAAATACAGGCAATTATTCATACAGCAACAAA 961
QY 21 ThwterSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerIeu 40
DB 962 ACTATGAGTGACACTTAACAGGCTCCCAAGCAAGAACTATTTATGACATCAAGTTTA 1021
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLeu 60
DB 1022 CCCACAGGCTCAGCTGGCTTCATATGAAATCTTTCTGACATTGATCAAGATGAAAACTC 1081
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1082 ACTGAGAGAAATTTATCTGATGATGACCTCATTTGATGCTATGTGCGATGCTGGTACGCA 1141
QY 81 LeuProValLeuProGlnTyrIleProSerPheArgArg 96
DB 1142 CTGCCGCCGCTGCTCCAGAAATACATTCCTCTCTTACAGAGA 1189
RESULT 12
AA180000
ID AA180000 standard; cDNA; 548 BP.
AC AA180000;
XX 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 60.
DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX Human; cytokine; cell proliferation; stem cell growth factor; haematopoiesis;
KM vaccine; peptide therapy; immunomodulatory; cancer; leukaemia;
KM tissue growth factor; immunomodulatory; cancer; leukaemia;
KM nervous system disorders; arthritis; inflammation; ss.
XX Homo sapiens.
OS
XX WO200164835-A2.
PN 07-SEP-2001.
XX 26-FEB-2001; 2001WO-US004927.
PF 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX (HSE-) HSEQ INC.
PA Tang YT, Liu C, Drmanac RT;
XX WPI: 2001-514838/56.
DR P-PSDB; AA000069.
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.

XX
PS Claim 1; SEQ ID NO 60; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 548 BP; 173 A; 124 C; 123 G; 127 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 9.77e-53 Length: 548
Score: 496.00 Matches: 95
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.00% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AA180000 (1-548)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 53 TGGGCTGTTCTCAGTATATACAGCTGAAATACAGGCAATATTCTAGTCACTGACAAA 112

QY 21 ThMeSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 113 ACTATAGTGTGACACTTAACAGGCCCCAAGCAAGAACTATTCTTATGAGTCAAGTTTA 172

QY 41 ProGlnAlaGlnLeuAlaSerIleTTPAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 173 CCACAGGCTCAGCTGGCTTAAATATGGAATCTTCTACACTTGATCAAGATGGAATACTT 232

QY 61 ThrAlaGlnIlePheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 233 ACAGCAGAGCAATTATCTCGCAATGCACTCATTTATGTAAGTCTGCGCAACCA 232

QY 81 LeuProProValLeuProProGlnTyrIleProProSerPheArg 95
Db 293 CTGCCACCTGTCTGCTCCAGATACATTCACCTCTTTTGA 337

RESULT 13
ABK43498
ID ABK43498 standard; cDNA; 3319 BP.
XX
AC ABK43498;
XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding novel central nervous system protein #78.
XX
XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
XX hyperproliferative disorder; neoplasm; cardiovascular disorder;
XX cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
XX nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
XX acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
XX adenocarcinoma; reproductive system disorder; testicular feminisation;
XX endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
XX respiratory disorder; renal disorder; kidney failure; blood disorder;
XX myocardial infarction; wound healing; cell proliferation; skin aging;
XX food additive; food preservative; gene therapy; gene; ss.
XX
XX Homo sapiens.
XX OS
XX PN MO20015518-A2.

XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001332.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-021513P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCT INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX

DR WPI; 2001-581633/65.
DR P-PSDB; JAU87168.
XX
XX
XX New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives or
PT preservatives.
XX
XX
PS Claim 1; SEQ ID NO 88; 837P; English.
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (II) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
Alignment Scores:
Pred. No.: 8.17e-51 Length: 3319
Score: 489.00 Matches: 96
Percent Similarity: 96.00% Conservative: 0
Best Local Similarity: 96.00% Mismatches: 0
Query Match: 97.60% Indels: 4
DB: Gaps: 1
US-09-720-934-2_COPY 215_310 (1-96) x ABK43498 (1-3319)
QY 1 TrrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 801 TGGGCTTCTCCTAGTATCAAGACTGAATACAGGCATTATTATCAATAGCATACAA 860
QY 21 ThrMetSerGlyHisLeuThr-----GlyProGlnAlaArgThrIleuMet 36
DB 861 ACTATGAGTGCACCTTAACAGGTTCTGTATTAGGTCCCAAGCAAGCAATCTTATG 920
QY 37 GlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGln 56
DB 921 CAGTCAGCTTTACACAGGCTCAGCTGCTTCATCAATATGAACTCTTCGACATTATCAA 980
QY 57 AspGlyLysLeuThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMet 76
DB 981 GATGGAAATCTTACAGCAGAGAAATTATCTGGCAATGACACCTGATGATGACTATG 1040
QY 77 SerGlyGlnProLeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
DB 1041 TCTGGCCAAACACCTGCGACCTGTCTGCTCCAGAAATCAATTCACCTTTTGAAGA 1100
RESULT 14
AA163825
ID AA163825 standard; cDNA, 3466 BP.
XX
XX AA163825;
AC
XX
XX 22-OCT-2001 (first entry)
DT
XX
XX Human polynucleotide SEQ ID NO 33.
DE
XX

PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM,
XX
XX WPI; 2001-488781/53.
XX P-PSDB; AAM43519.
XX
XX New isolated nucleic acids and polypeptides, useful for diagnosing,
XX treating and/or preventing human diseases and disorders.
XX
XX Claim 1; SEQ ID NO 33; 664bp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA163803-AA164012) and
XX the encoded proteins (AAM434497-AAM43660) useful for preventing, treating
XX or ameliorating medical conditions e.g. by protein or gene therapy. The
XX genes were isolated from a range of human tissues disclosed in the
XX specification. The nucleic acids, proteins, antibodies and (ant)agonists
XX are useful in the diagnosis, treatment and prevention of: (a) cancer,
XX e.g. breast and ovarian cancer and other cancers of the adrenal gland,
XX b) immune disorders e.g. Addison's disease, allergies,
XX c) autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
XX Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)
XX wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections. Note: The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 3466 BP; 1056 A; 829 C; 861 G; 710 T; 0 U; 10 Other;
XX
XX Alignment Scores:
XX Pred. No.: 8.65e-51 Length: 3466
XX Score: 489.00 Matches: 96
XX Percent Similarity: 96.00% Conservative: 0
XX Best Local Similarity: 96.00% Mismatches: 0
XX Query Match: 97.60% Indels: 4
XX DB: Gaps: 1
XX
XX US-09-720-934-2_COPY_215_310 (1-96) x AA163825 (1-3466)
XX
XX 1 TTPALAVAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
XX 919 TGGGTGTTCTCCATCATCAAGACTGAAATACAGGCAATATTTCAATAGTCATACAAA 978
XX 21 ThrMetSerGlyHisLeuThr-----GlyProGlnAlaArgThrIleLeuMet 36
XX 979 ACTATGAGTGCACACTTAACAGGTTCCGTTAGGTCCCAAGCAAGAATCTTCTTAG 1038
XX 37 GlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTyrPheAsnSerHisAspLys 56
XX 1039 CAGTCAGATTACCAACAGGCTGAGCTGCTTCAAAATGAAATCTTTCGACATTCATCA 1098
XX 57 AspGlyLysLeuThrAlaGlnIlePheIleLeuAlaMetHisLeuIleAspValAlaMet 76
XX 1099 GATGGAAAACTTACGACAGAGAAATTTATCTCGGCAATACACCTCATTCATTAAGTANG 1158
XX 77 SerGlyGlnProLeuProProValLeuProProGlnTyrIleProProSerPheArgArg 96
XX 1159 TCTGGCCAAACCACTCCGCTCTCTGCTCCAGATATACATTCACCTTCTTTAGAGA 1218
XX
XX RESULT 15
XX AAS84763
XX ID AAS84763 standard; cDNA; 7435 BP.
XX

AC AAS84763;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #20567.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX P-PSDB; ABG20576.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 20567; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3.44e-41 Length: 7435
XX Score: 416.00 Matches: 96
XX Percent Similarity: 93.20% Conservative: 0
XX Best Local Similarity: 93.20% Mismatches: 0
XX Query Match: 83.03% Indels: 7
XX DB: Gaps: 0
XX
XX US-09-720-934-2_COPY_215_310 (1-96) x AAS84763 (1-7435)
XX
XX 1 TTPALAVAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAsp 19
XX 907 TGGGTGTTCTCCATCATCAAGACTGAAATATACAGGCAATATTTCAATAGTCATGAC 966
XX 20 LysThr--MetSerGlyHis-LeuThr-GlyProGlnAlaArgThrIleLeuMetClnse 38
XX

Db 967 ||||| 1026
QY 38 rSerleupProGlnAgiInleuAlaSerIleTnpAsnleuSerAspIleAspGlnAspG1 58
Db 1027 AAGTTTACCACAGGCTCAGCTGCTTCAATATGGAATCTTCTGACATTGATCAAGATGG 1086
QY 58 YLysleuThrAlaGlnGluPheIleLeuAlaMetHisleuIleAspValAlaMetSerg1 78
Db 1087 AAACTTACACAGCAGGAAATTATCTGCGCAATGCACTCATTTGATGCTATGCTCG 1146
QY 78 YGlnProleupProProValleuProProGlnuTyrlleProProSerPheArgArg 96
Db 1147 CCAACCACTGCCACCTGTCTGCTCCGAAATACATTCCACCTTCTTTTGAAGA 1201

Search completed: July 1, 2004, 13:31:36
Job time : 305.319 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 67.2225 Seconds
(without alignments)
792,522 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 WAPVQSSRLKRLQLENSHDK.....SGQPLPPVLPPEYTPPSFRR 96

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/p/US09720934/runat.30062004.064540.1345/app.query.fasta_1.1386
-DB=Issued Patents NA -OEMT=fastlap -SUFFIX=rni -MINMATCH=0.1 -LOOPTC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720934 @CGN 1 1 284 @runat.30062004.064540.13455 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2.6/prodata/2/ina/5A COMB.seg:*
2: /cgn2.6/prodata/2/ina/5B COMB.seg:*
3: /cgn2.6/prodata/2/ina/6A COMB.seg:*
4: /cgn2.6/prodata/2/ina/6B COMB.seg:*
5: /cgn2.6/prodata/2/ina/FACTUS COMB.seg:*
6: /cgn2.6/prodata/2/ina/backfile1.seg:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	69.7	2017	4	US-09-404-879A-72
2	349	69.7	2017	4	US-09-358-933-72
3	349	69.7	2017	4	US-09-215-681-72
4	349	69.7	2017	4	US-09-216-003A-72
5	208	41.5	480	4	US-09-404-879A-60
6	208	41.5	480	4	US-09-338-933-60
7	208	41.5	480	4	US-09-215-681-60
8	208	41.5	480	4	US-09-216-003A-60
9	208	41.5	531	4	US-09-404-879A-5
10	208	41.5	531	4	US-09-338-933-5
11	208	41.5	531	4	US-09-215-681-5
12	208	41.5	531	4	US-09-216-003A-5

13	204.5	40.8	3033	1	US-08-095-737-3	Sequence 3, Appl1
14	204.5	40.8	3033	2	US-08-480-145-3	Sequence 3, Appl1
15	204.5	40.8	3033	1	US-08-477-389-3	Sequence 3, Appl1
16	197.5	39.4	4165	1	US-08-095-737-1	Sequence 1, Appl1
17	197.5	39.4	4165	1	US-08-480-145-1	Sequence 1, Appl1
18	197.5	39.4	4165	2	US-08-477-389-1	Sequence 2, Appl1
19	189	37.7	3348	4	US-09-312-762A-2	Sequence 1, Appl1
20	189	37.7	3508	4	US-09-312-762A-1	Sequence 1, Appl1
21	189	37.7	14707	4	US-09-312-762A-3	Sequence 3, Appl1
22	176	35.1	3635	4	US-09-312-762A-6	Sequence 6, Appl1
23	170	33.9	2373	4	US-09-023-655-850	Sequence 850, App
24	162	32.3	590	3	US-09-328-111-211	Sequence 211, App
25	151	30.1	2815	4	US-09-312-762A-7	Sequence 7, Appl1
26	138	27.5	5764	4	US-09-312-762A-10	Sequence 10, Appl1
27	94.5	18.9	845	3	US-08-998-416-543	Sequence 543, App
28	94	18.8	503	4	US-09-370-838-148	Sequence 148, App
29	90.5	18.1	884	4	US-09-833-381-1048	Sequence 1048, App
30	87	17.4	853	3	US-08-963-409-2	Sequence 2, Appl1
31	81.5	16.3	1776	3	US-08-655-352-10	Sequence 10, Appl1
32	81.5	16.3	1776	4	US-09-258-016-10	Sequence 10, Appl1
33	81.5	16.3	1776	4	US-09-257-825B-10	Sequence 8, Appl1
34	81	16.2	1201	3	US-09-048-889-8	Sequence 8, Appl1
35	81	16.2	2259	4	US-09-399-913-71	Sequence 69, Appl1
36	81	16.2	2259	4	US-09-399-913-69	Sequence 48, Appl1
37	81	16.2	2413	4	US-09-399-913-48	Sequence 50, Appl1
38	79	15.8	1591	4	US-09-338-671-3	Sequence 31, Appl1
39	79	15.8	1835	3	US-09-399-913-31	Sequence 35, Appl1
40	78	15.6	2841	4	US-09-399-913-35	Sequence 35, Appl1
41	76	15.2	2644	4	US-09-298-731-35	Sequence 5, Appl1
42	76	15.2	2644	2	US-08-818-252-5	Sequence 5, Appl1
43	75	15.0	1929	3	US-08-818-252-5	Sequence 7, Appl1
44	75	15.0	1929	2	US-08-818-253-7	Sequence 7, Appl1
45	75	15.0	1971	2	US-08-818-253-7	Sequence 7, Appl1

ALIGNMENTS

```
RESULT 1
US-09-404-879A-72
; Sequence 72, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404, 879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-72

Alignment Scores:
Pred. No.: 1,35e-37
Score: 349.00
Percent Similarity: 82.11%
Best Local Similarity: 71.58%
Query Match: 69.66%
DB: 4
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-404-879A-72 (1-2017)
QY 1 TYPAlaValProGlnSerSerArgLeuLyTYArgGlnLeuPheAsnSerHisAspLys 20
Db 753 TGGGCGAGTTCCTCAGCCCTTCAGATTAAGATTCACGCAAAATTTATACCTCAGACAAA 812
QY 21 ThMeSerGlyHisLeuThrGlyProGlnAlaArgThrIleuMetGlnSerSerLeu 40
```

```
Db 813 GGCATGACGGATACCTCTTCAAGCTAATAATGCCCTTCTTCAGTCAAAATCTC 872
Qy 41 ProglinalaginLeuA1Ser1eTTPasnLeuSerAp1leASPglAspGlyLysleu 60
Db 873 TCTCAAACTCAGCTAGCTACTATTGAGCTCTGGCTGACATCGATGAGTGCAGCAGATTG 932
Qy 61 ThrAlaglugluPhe1leLeuA1AmetHisLeu1leaspVal1AlaMetSerGlyGlnPro 80
Db 933 AAAGCTGAAGAATTATTCTTGGCGATGCACTCCTCACTGACATGCGCAAAAGCTGGACAGCCA 992
Qy 81 LeuProProVal1leuProProGluTyr1leProProSerPheArg 95
Db 993 CTACCACTGACGTTGCTCCGAGCTTGCTCCCTCATCTTTTCAGA 1037

RESULT 2
US-09-338-933-72
; Sequence 72, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-72

Alignment Scores:
Pred. No.: 1,35e-37 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-338-933-72 (1-2017)
Qy 1 TTPAlava1ProglInSerSerArgleuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 753 TGGGAGATTCTCTCAGCCTTCAAGATTAAAGTATCGGCAAAAATTAAATAGTCTAGACAAA 812
Qy 21 ThmetSerGlyHisLeuThnGlyProGln1AaTgThr1leLeuMetGlnSerSerleu 40
Db 813 GGCATGACGGATACCTCTCAGCTTTCAAGCTAATAATGCCCTTCTTCAGTCAAAATCTC 872
Qy 41 ProglinalaginLeuA1Ser1eTTPasnLeuSerAp1leASPglAspGlyLysleu 60
Db 873 TCTCAAACTCAGCTAGCTACTATTGAGCTCTGGCTGACATCGATGAGTGCAGCAGATTG 932
Qy 61 ThrAlaglugluPhe1leLeuA1AmetHisLeu1leaspVal1AlaMetSerGlyGlnPro 80
Db 933 AAAGCTGAAGAATTATTCTTGGCGATGCACTCCTCACTGACATGCGCAAAAGCTGGACAGCCA 992
Qy 81 LeuProProVal1leuProProGluTyr1leProProSerPheArg 95
Db 993 CTACCACTGACGTTGCTCCGAGCTTGCTCCCTCATCTTTTCAGA 1037

RESULT 3
US-09-215-681-72
; Sequence 72, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-72

Alignment Scores:
Pred. No.: 1,35e-37 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-215-681-72 (1-2017)
Qy 1 TTPAlava1ProglInSerSerArgleuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 753 TGGGAGATTCTCTCAGCCTTCAAGATTAAAGTATCGGCAAAAATTAAATAGTCTAGACAAA 812
Qy 21 ThmetSerGlyHisLeuThnGlyProGln1AaTgThr1leLeuMetGlnSerSerleu 40
Db 813 GGCATGACGGATACCTCTCAGCTTTCAAGCTAATAATGCCCTTCTTCAGTCAAAATCTC 872
Qy 41 ProglinalaginLeuA1Ser1eTTPasnLeuSerAp1leASPglAspGlyLysleu 60
Db 873 TCTCAAACTCAGCTAGCTACTATTGAGCTCTGGCTGACATCGATGAGTGCAGCAGATTG 932
Qy 61 ThrAlaglugluPhe1leLeuA1AmetHisLeu1leaspVal1AlaMetSerGlyGlnPro 80
Db 933 AAAGCTGAAGAATTATTCTTGGCGATGCACTCCTCACTGACATGCGCAAAAGCTGGACAGCCA 992
Qy 81 LeuProProVal1leuProProGluTyr1leProProSerPheArg 95
Db 993 CTACCACTGACGTTGCTCCGAGCTTGCTCCCTCATCTTTTCAGA 1037

RESULT 4
US-09-216-003A-72
; Sequence 72, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-72

Alignment Scores:
Pred. No.: 1,35e-37 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-216-003A-72 (1-2017)
```

```

OY      1 TTPALAVAIProGlnSerSerSerArgLeuValSTYArgGlnLeuPheAsnSerHisAspGlys 20
Db      753 TGGGCAAGTTCCTCAGCCTTCAGATTAAAGTATCGGCAGAAAATTTAATAGCTAGACAAA 8122
OY      21 ThrMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db      813 GGCATAGAGCGGATACCTCTCAGCGTTTTCAGAGTAAAGTACCCCTCTTCAGTCAAAATCTC 872
OY      41 ProGlnAlaGlnLeuAlaSerIleTPanLeuSerAspIleAspGlnAspGlyLeu 60
Db      873 TCTCAAACTCAGCTACCTACTACTATTGTAAGCTTGAGCTGACATGAGTGTGACGACAGTTG 922
OY      61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db      933 AAAGCTGAAGATTATTTCTGGGAGTAGCAACCTCAGTACGATGCGCAAAAGCTGACAGCCA 992
OY      81 LeuProProValLeuProProGlnGlyrIleProProSerPheArg 95
Db      993 CTACCACTGACGTTGCTCCCGAGCGCTTGTCCTCCACACTTTACGA 1037

```


Alignment Scores:	
Pred. No.:	3,08e-19
Score:	208.00
Percent Similarity:	63.16%
Best Local Similarity:	46.32%
Query Match:	41.52%
Ds:	4
Length:	531
Matches:	44
Conservative:	16
Mismatches:	33
Indels:	
Gaps:	2

US-09-720-934-2_COPY_215_310 (1-96) X US-09-338-933-5 (1-531)

QY 1 TPPLAVALAProGlnSerSerArgLeuValysArgGlnLeuPheAnsSerHisAspLys 20
 Db 84 TGGGCTATTACATCTGAAGAAGCCTACTAGACATGATTAACAGTTTGATTAACCTC--AAA 14
 QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaIArgThrIleLeuMetGlnSerSerLeu 40
 Db 141 CCTTGAGGAGTTACTTAACAGAGGTGATCAAGCCCGTACTTTTCTTACACGTAGAGCTTG 200
 QY 41 ProGlnAlaGlnLeuAlaSerIleTETPAuLeuSerAspIleAspGlnAspGlyLysLeu 60
 Db 201 CCGGCCCCGGTTTAACTGAAATATATGGGCTTTATCGATGTGAACAAGATGGGAAATATG 260
 QY 61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 Db 261 GACCAACGAAGAGTTCTCTATAGCTATGAAACATCATCAAGATTAAAGTTGACAGGCCAACAG 320
 QY 81 LeuProProValLeuProPro--GluTyrIleProProSerPhe 94
 Db 321 CTGCGCTGTAAGTCTCTCCCTCATATGAAACAACCCCTCATATGTTTC 365

RESULT 11

```

US-09-215-681-5
; Sequence 5, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF OVARIAN CANCER
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-5

```

Alignment Scores:	
Pred. No.:	3.08e-19
Score:	208.00
Percent Similarity:	63.16%
Best Local Similarity:	46.32%
Query Match:	41.52%
DB:	4
Gaps:	2
length:	531
Matches:	44
Conservative:	16
Mismatches:	33
Indels:	3
Gaps:	2

US-09-720-934-2_COPY_215_310 (1-96) x US-09-215-681-5 (1-531)

```

QY 1 TTPALAVAlProGlnIseSerIArgLeuYTYArgGlnIleuPheAnsSerHisAspLys 20
      ::::: ::::: ::::: ::::: :::::
Db 84 TGGGCTATTACACTGTGAGAACGCTACTAGACATGATATAACAGTTGTATACCTC---AAA 140
QY 21 ThMeSeSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnIseSerLeu 40
      ::::: ::::: ::::: ::::: :::::
Db 141 CCTCAGGAGGTTACATACAGGTGATCAAGCCGAGCTTTTTCCTACAGCTCAGGCTCG 200
QY 41 ProGlnAlaGlnLeuAlaSerIleTPaMLeuSerAspIleAspGlnAspGlyLeu 60
      ::::: ::::: ::::: ::::: :::::
Db 201 CCGGCCCCGGATTTCCTGAAAATATGGCCCTTTCAGATTCATGAACAAGATGGGAATAG 260

```

[illegible]

RESULT 12

```

US-09-216-003A-5
; Sequence 5, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-5

```

Pred. No.:	3.08e-19	Length:	531
Score:	208.00	Matches:	44
Percent Similarity:	63.16%	Conservative:	16
Best Local Similarity:	46.32%	Mismatches:	33
Query Match:	41.52%	Indels:	2
DB:	4	Gaps:	2

US-09-720-934-2_COPY_215_310 (1-96) x US-09-216-003A-5 (1-531)

[illegible]

RESULT 13

US-08-095-137-3
 ? Sequence 3, Application US/08095737
 ? Patent No. 5487979
 ?
 ? GENERAL INFORMATION:
 ?
 ? APPLICANT: DiFiore, Pier P
 ? APPLICANT: Fazioli, Francesca
 ?
 ? TITLE OF INVENTION: A Substrate for the Epidermal G
 ?
 ? TITLE OF INVENTION: Factor Receptor Kinase
 ?
 ? NUMBER OF SEQUENCES: 4
 ?
 ? CORRESPONDENCE ADDRESS:
 ?
 ? ADDRESSEE: Knobbe, Martens, Olson & Bear
 ?
 ? STREET: 620 Newport Center Drive, Sixteenth Floor
 ?
 ? CITY: Newport Beach

STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,737
FILING DATE: 19930722
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-095-737-3

Alignment Scores:
Pred. No.: 1.14e-17 Length: 3033
Score: 204.50 Matches: 40
Percent Similarity: 60.42% Conservative: 18
Best Local Similarity: 41.67% Mismatches: 37
Query Match: 40.82% Indels: 1
DB: 1 Gaps: 1

US-09-720-934-2_COPY_215_310 (1-96) x US-08-095-737-3 (1-3033)
QY 1 ThrpAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 474 TGGGCTGTAAGTCTGAAGATPAAAGCCAAATATGATGCATTTTGAACAGTTTAA--AGC 530
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 531 CCAGTGGATGATTTTGTCTGCGATPAAAGTGAACAGGTTGCTCACTCACTTAAGTTA 590
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
DB 591 CCTGTGAAATCCTTGGAAAGATTGGAGTTGAGTGAATTATTTGACACAGATGGAAGCTG 650
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 651 GACAGAGATGAGTTTGCAGTTGCCATGTTTGTGATATCTGTGACCTGGAAGAAAGAACTT 710
QY 81 LeuProValLeuProGluTyrIleProProSerPheArg 96
DB 711 GTGCCAATGTCCTTGCCTCCAGCCTTGGTGTCACCTTCTAAGAGAAA 758

RESULT 14
US-08-480-145-3
Sequence 3, Application US/08480145
Patent No. 5717067
GENERAL INFORMATION:
APPLICANT: DiFiore, Pier P
APPLICANT: Razioli, Francesca
TITLE OF INVENTION: A Substrate for the Epidermal Growth
Factor Receptor Kinase
NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,145
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/095,737
FILING DATE: 22-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-480-145-3

Alignment Scores:
Pred. No.: 1.14e-17 Length: 3033
Score: 204.50 Matches: 40
Percent Similarity: 60.42% Conservative: 18
Best Local Similarity: 41.67% Mismatches: 37
Query Match: 40.82% Indels: 1
DB: 1 Gaps: 1

US-09-720-934-2_COPY_215_310 (1-96) x US-08-480-145-3 (1-3033)
QY 1 ThrpAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 474 TGGGCTGTAAGTCTGAAGATPAAAGCCAAATATGATGCATTTTGAACAGTTTAA--AGC 530
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 531 CCAGTGGATGATTTTGTCTGCGATPAAAGTGAACAGGTTGCTCACTCACTTAAGTTA 590
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
DB 591 CCTGTGAAATCCTTGGAAAGATTGGAGTTGAGTGAATTATTTGACACAGATGGAAGCTG 650
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 651 GACAGAGATGAGTTTGCAGTTGCCATGTTTGTGATATCTGTGACCTGGAAGAAAGAACTT 710
QY 81 LeuProValLeuProGluTyrIleProProSerPheArg 96
DB 711 GTGCCAATGTCCTTGCCTCCAGCCTTGGTGTCACCTTCTAAGAGAAA 758

RESULT 15
US-08-477-389-3
Sequence 3, Application US/08477389

Patent No. 5872219
GENERAL INFORMATION:
APPLICANT: DiFiore, Pier P
APPLICANT: Fazio, Francesca
TITLE OF INVENTION: A Substrate for the Epidermal Growth
FACTOR RECEPTOR KINASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,389
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/095,737
FILING DATE: 22-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-477-389-3

Alignment Scores:

Pred. No.:	1,14e-17	Length:	3033
Score:	204.50	Matches:	40
Percent Similarity:	60.42%	Conservative:	18
Best Local Similarity:	41.67%	Mismatches:	37
Query Match:	40.82%	Indels:	1
DB:	2	Gaps:	1

US-09-720-934-2_COPY_215_310 (1-96) x US-08-477-389-3 (1-3033)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 474 TGGGCTGTAAAGCTGAAGTAAAGCCAAATATGATGCAATTTTGAACAGTTTA--AGC 530
QY 21 ThMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 531 CCAGTGATGATGATTTTGTGTGTGATTAAGTGAACCAAGTTGCTCAACTCTAAGTTA 590
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 591 CCTGTGGAATCTCTTGAAGAGTTTGGAGTTGAGTATTTGACCAAGTGAAGCTG 650
QY 61 ThrAlaGlnGlnPheIleuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 651 GACAGAGATGAGTTTGCAGTGTCCATGTTTGTATACTGTGCACTGAGAAAGAACT 710

QY 81 LeuProProValLeuProProGlnUtyrIleProProSerPheArgArg 96
Db 711 GTGCCAATGTCCTTGCTCCAGGCTTGGTGCCACCTTTCAAGAGAAA 758
Search completed: July 1, 2004, 19:53:41
Job time : 71.2225 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 294.52 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 MAVPQSSRLKXRLPFSHDK.....SQGPLPVLPPEITPSPFR 96

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O/cgn2_1/USFTO.spool_p/US09720934/runat_30062004_064541_13481/app_query_faasta_1.1386
-DB=Published Applications NA -OPMT=fastrep -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALL=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADING=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09720934.@CGN_1_1500.@runat_30062004_064541_13481
-NCPU=6 -ICPU=3 -NO MAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : Published Applications_NA:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query SUMMARIES
No. Score Match Length DB ID Description

1	489	97.6	3319	11	US-09-764-875-88	Sequence 88, Appl
2	489	97.6	3466	16	US-10-158-057-33	Sequence 33, Appl
3	349	69.7	2017	9	US-09-884-441-72	Sequence 72, Appl
4	349	69.7	2017	10	US-09-907-969-72	Sequence 72, Appl
5	349	69.7	2017	10	US-09-827-271-72	Sequence 72, Appl
6	349	69.7	2017	15	US-10-198-053-72	Sequence 72, Appl
7	345	68.9	5828	13	US-10-398-885A-15	Sequence 15, Appl
8	334	66.7	955	10	US-09-764-881-54	Sequence 54, Appl
9	334	66.7	955	11	US-09-764-875-91	Sequence 91, Appl
10	334	66.7	955	16	US-10-242-747-54	Sequence 54, Appl
11	334	66.7	955	16	US-10-242-747-54	Sequence 54, Appl
12	334	66.7	955	16	US-10-158-057-33	Sequence 34, Appl
13	218.5	43.6	3976	16	US-10-369-493-36727	Sequence 36727, A
14	208	41.5	480	9	US-09-884-441-60	Sequence 60, Appl
15	208	41.5	480	10	US-09-907-969-60	Sequence 60, Appl
16	208	41.5	480	10	US-09-827-271-60	Sequence 60, Appl
17	208	41.5	480	15	US-10-198-053-60	Sequence 60, Appl
18	208	41.5	531	9	US-09-884-441-5	Sequence 5, Appl
19	208	41.5	531	10	US-09-907-969-5	Sequence 5, Appl
20	208	41.5	531	10	US-09-827-271-5	Sequence 5, Appl
21	208	41.5	531	15	US-10-198-053-5	Sequence 5, Appl
22	191	38.1	2259	16	US-10-369-493-26020	Sequence 26020, A
23	189	37.7	462	10	US-09-918-995-21728	Sequence 21728, A
24	189	37.7	605	12	US-10-152-319A-1126	Sequence 1126, Ap
25	189	37.7	3348	9	US-09-312-762A-2	Sequence 2, Appl
26	189	37.7	3508	9	US-09-312-762A-1	Sequence 1, Appl
27	189	37.7	14707	9	US-10-027-632-165796	Sequence 165796,
28	188	37.5	860	13	US-10-027-632-165796	Sequence 165796,
29	188	37.5	860	13	US-10-027-632-165797	Sequence 165797,
30	188	37.5	860	16	US-10-027-632-165796	Sequence 165796,
31	188	37.5	860	16	US-10-027-632-165797	Sequence 165797,
32	188	37.5	3485	15	US-10-171-581-354	Sequence 354, App
33	187.5	37.4	2442	9	US-09-964-899-24	Sequence 24, Appl
34	187	37.3	662	13	US-10-027-632-290824	Sequence 290824,
35	187	37.3	662	16	US-10-027-632-290824	Sequence 290824,
36	185	36.9	523	10	US-09-814-353-14688	Sequence 14688, A
37	185	36.9	707	10	US-09-764-881-186	Sequence 186, App
38	185	36.9	707	13	US-09-764-881-186	Sequence 186, App
39	185	36.9	707	16	US-10-242-747-186	Sequence 186, App
40	185	36.9	707	16	US-10-158-057-373	Sequence 373, App
41	185	36.9	2105	11	US-09-764-875-175	Sequence 175, App
42	185	36.9	2414	10	US-09-814-353-20215	Sequence 20215, A
43	185	36.9	2907	16	US-10-108-260A-1761	Sequence 1761, Ap
44	185	36.9	3017	9	US-09-818-143-8	Sequence 8, Appl
45	184	36.7	523	10	US-09-814-353-1957	Sequence 1957, Ap

ALIGNMENTS

RESULT 1
US-09-764-875-88
; Sequence 88, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 3319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-875-88

Alignment Scores: 8.22e-60 Length: 3319
Pred. No.: 489.00 Matches: 96
Score: 96.00% Conservative: 0

Best Local Similarity: 96.00%
Query Match: 97.60%
DB: 11
Gaps: 1

US-09-720-934-2_COPY_215_310 (1-96) x US-09-764-875-88 (1-3319)

QY 1 TrpAlaValProGlnSerSerArgLeuYsTYrArgGlnLeuPheAsnSerHisAspLys 20
DB 801 TGGGCTGTTCCTCAGTATCAAGACTGAATACAGGCAATTATTAATAGTCATACAA 860

QY 21 ThrMetSerGlyHisLeuThr-----GlyProGlnAlaArgThrIleLeuMet 36
DB 861 ACTAGAGCTGACCACTTAACTGCTTCTGTTAGGCTCCCAAGCAACTATTCTTATG 920

QY 37 GlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGln 56
DB 921 CAGTCAGTTTACCAAGAGGCTGCTGCTGCTTCAATATGAACTTTTGACATTTGATCAA 980

QY 57 AspGlyLysLeuThrAlaGluGlnPheIleLeuAlaMetHisLeuIleAspValAlaMet 76
DB 981 GATGGAAACTTACAGCAGAGGATTTATCTGGCAATGCACCTCATTTGATGACTATG 1040

QY 77 SerGlyGlnProLeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
DB 1041 TCTGGCCACCACTGCGCACTGCTCTGCTCCAGAAATCATTTCCACTTTTGAAGA 1100

RESULT 2
US-10-158-057-33
; Sequence 33, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P120501
; CURRENT APPLICATION NUMBER: US/10/158,057
; CURRENT FILING DATE: 2002-06-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 3466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3194)
; OTHER INFORMATION: n equals a,t,b, or c
; NAME/KEY: misc feature
; LOCATION: (3465)
; OTHER INFORMATION: n equals a,t,b, or c
US-10-158-057-33

Alignment Scores:
Pred. No.: 8.73e-60
Score: 489.00
Percent Similarity: 96.00%
Best Local Similarity: 96.00%
Query Match: 97.60%
DB: 16
Length: 3466
Matches: 96
Conservative: 0
Mismatch: 0
Indels: 4
Gaps: 1

US-09-720-934-2_COPY_215_310 (1-96) x US-10-158-057-33 (1-3466)

QY 1 TrpAlaValProGlnSerSerArgLeuYsTYrArgGlnLeuPheAsnSerHisAspLys 20
DB 919 TGGGCTGTTCCTCAGTATCAAGACTGAATACAGGCAATTATTAATAGTCATACAA 978

QY 21 ThrMetSerGlyHisLeuThr-----GlyProGlnAlaArgThrIleLeuMet 36
DB 979 ACTATGAGTGACACTTAACAGCTTCTGTTAGTCCCAAGCAAGACTATTCTTATG 1038

QY 37 GlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGln 56

DB 1039 CAGTCAAGTTTACACAGAGCTGAGCTTCATATAGAAATCTTTCGACATTGATCAA 1098

QY 57 AspGlyLysLeuThrAlaGluGlnPheIleLeuAlaMetHisLeuIleAspValAlaMet 76
DB 1099 GATGGAAACTTACAGCAGAGGAAATTTATCTGGCAATGCACCTCATTTGATGACTATG 1158

QY 77 SerGlyGlnProLeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
DB 1159 TCTGGCCACCACTGCGCACTGCTCTGCTCCAGAAATCATTTCCACTTTTGAAGA 1218

RESULT 3
US-09-884-441-72
; Sequence 72, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darlick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-72

Alignment Scores:
Pred. No.: 7.22e-40
Score: 349.00
Percent Similarity: 82.11%
Best Local Similarity: 71.58%
Query Match: 69.66%
DB: 9
Length: 2017
Matches: 68
Conservative: 10
Mismatch: 17
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-884-441-72 (1-2017)

QY 1 TrpAlaValProGlnSerSerArgLeuYsTYrArgGlnLeuPheAsnSerHisAspLys 20
DB 753 TGGGCTGTTCCTCAGTATCAAGACTGAATTAAGATGCGCAAAATTTAATAGTCATACAA 812

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 813 GGCATGAGCGGATACCTCTCAGGTTTCAAGCTAGAAATGCCCTTTCAGTCAAAATCTC 872

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheAsnSerAspIleAspGlnAspGlyLysLeu 60
DB 873 TCTGAACCTCAGCTAGCTACATTTTGACTCTGCTGACATCGATGTCGACAGCAGTTG 932

QY 61 ThrAlaGluGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 933 AAAGCTGAGAAATTTATTTCTGCGATGACACTCACTACATGCGCAAGCTGACAGCCA 992

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
DB 993 CTACCACTGACGTTGCTCCGAGCTTGTCTCTCATTTTCAGA 1037

RESULT 4
US-09-907-969-72
; Sequence 72, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Ranger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.

```
; APPLICANT: Carter, Darrick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-907-969-72

Alignment Scores:
Pred. No.: 7,22e-40 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-907-969-72 (1-2017)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
   |||||
Db 753 TGGGCAgTTCCTCAGCCTTCAGATTAAAGTATCGGCAAAATTTAATAGCTAGACAAA 812

QY 21 ThrMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
   |||||
Db 813 GGCATAGCGGATACCTCTCAGGATCGACCTGCAATGCGCAAAAGCTGGACAGCCA 992

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysLeu 60
   |||||
Db 873 TCTCAAACTCAGCTAGTACTATTGGAAGCTGCTGCTGACATGCTGACAGCAGTTG 932

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
   |||||
Db 933 AAAGCTGAAGAATTATTTCTGGCGATGCACTGACCTGACCTGCAAAAGCTGGACAGCCA 992

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
   |||||
Db 993 CTACCACTGACGTTGCTCCGAGCTTGTCTTCATCTTTCAGA 1037

RESULT 5
US-09-827-271-72
; Sequence 72, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C6
; CURRENT APPLICATION NUMBER: US/09/827,271
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-827-271-72

Alignment Scores:
Pred. No.: 7,22e-40 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: 10 Gaps: 0
```

```
US-09-720-934-2_COPY_215_310 (1-96) x US-09-827-271-72 (1-2017)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
   |||||
Db 753 TGGGCAgTTCCTCAGCCTTCAGATTAAAGTATCGGCAAAATTTAATAGCTAGACAAA 812

QY 21 ThrMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
   |||||
Db 813 GGCATAGCGGATACCTCTCAGGATCGACCTGCAATGCGCAAAAGCTGGACAGCCA 992

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysLeu 60
   |||||
Db 873 TCTCAAACTCAGCTAGTACTATTGGAAGCTGCTGCTGACATGCTGACAGCAGTTG 932

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
   |||||
Db 933 AAAGCTGAAGAATTATTTCTGGCGATGCACTGACCTGACCTGCAAAAGCTGGACAGCCA 992

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
   |||||
Db 993 CTACCACTGACGTTGCTCCGAGCTTGTCTTCATCTTTCAGA 1037

RESULT 6
US-10-198-053-72
; Sequence 72, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-198-053-72

Alignment Scores:
Pred. No.: 7,22e-40 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: 15 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-10-198-053-72 (1-2017)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
   |||||
Db 753 TGGGCAgTTCCTCAGCCTTCAGATTAAAGTATCGGCAAAATTTAATAGCTAGACAAA 812

QY 21 ThrMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
   |||||
Db 813 GGCATAGCGGATACCTCTCAGGATCGACCTGCAATGCGCAAAAGCTGGACAGCCA 992

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysLeu 60
   |||||
Db 873 TCTCAAACTCAGCTAGTACTATTGGAAGCTGCTGCTGACATGCTGACAGCAGTTG 932

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
   |||||
Db 933 AAAGCTGAAGAATTATTTCTGGCGATGCACTGACCTGACCTGCAAAAGCTGGACAGCCA 992

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
   |||||
Db 993 CTACCACTGACGTTGCTCCGAGCTTGTCTTCATCTTTCAGA 1037
```

Db 993 CTACCACTGACGTTGCCGAGCTGTCTCCCTCCATCTTTGAGA 1037

RESULT 7
US-10-398-885A-15
; Sequence 15, Application US/10398885A
; Publication No. US20040053282A1
; GENERAL INFORMATION:
; APPLICANT: Sugita, Yuji
; APPLICANT: Hashida, Ryoichi
; APPLICANT: Ogawa, Kaoru
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Obayashi, Masaya
; APPLICANT: Saito, Hirohisa
; APPLICANT: Takahashi, Eiki
; TITLE OF INVENTION: Method of Testing For Allergic Diseases
; FILE REFERENCE: SHIMIZU-07907
; CURRENT APPLICATION NUMBER: US/10/398, 885A
; PRIORITY FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: PCT/JP01/08937
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: JP 2000-314093
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 5828
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: CDS
; LOCATION: (7)..(5052)
; OTHER INFORMATION:
US-10-398-885A-15

Alignment Scores:
Pred. No.: 1.19e-38 Length: 5828
Score: 345.00 Matches: 65
Percent Similarity: 83.16% Conservative: 14
Best Local Similarity: 68.42% Mismatches: 16
Query Match: 68.86% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-10-398-885A-15 (1-5828)

QY 1 TTPAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspIys 20
Db 751 TGGGACAGTCTCCTCAGCCTACAAAGATTAAATATCGGCAAAATTTAATACCTTGACAAA 810
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
Db 811 AGTATGATGATATATCTCTCAGGTTTTCAGCTAGAAATGCCCTTCTTCAGTCAAAATCTT 870
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
Db 871 TCTCAACTCAGCTGGCTGACTACTTGGACTCTGCTGACGTTGATGATGATGACAGCTA 930
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 931 AAACGAGAGAGGTTTATCTTGCAATGACCTTACTGACATGCGCCAAAGCTGACAGCCA 990
QY 81 LeuProProValLeuProProGlnTyrIleProProSerPheArg 95
Db 991 TTACCACTGACTTACCTCCTGAGCTTGTCTCCATCTTTGAGA 1035

RESULT 8
US-09-764-881-54
; Sequence 54, Application US/0976481
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - refer to PAM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (550)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (934)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (942)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (944)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (948)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-54

Alignment Scores:
Pred. No.: 3.78e-38 Length: 955
Score: 334.00 Matches: 64
Percent Similarity: 81.05% Conservative: 13
Best Local Similarity: 67.37% Mismatches: 18
Query Match: 66.67% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-764-881-54 (1-955)

QY 1 TTPAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspIys 20
Db 390 TGGGACAGTCTCCTCAGCCTACAAAGATTAAATATCGGCAAAATTTAATACCTTGACAAA 449
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
Db 450 AGTATGATGATATATCTCTCAGGTTTTCAGCTAGAAATGCCCTTCTTCAGTCAAAATCTT 509
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
Db 510 TCTCAACTCAGCTGGCTGACTACTTGGACTCTGCTGACGTTGATGATGATGACAGCTA 569
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 570 AAACGAGAGAGGTTTATCTTGCAATGACCTTACTGACATGCGCCAAAGCTGACAGCCA 629
QY 81 LeuProProValLeuProProGlnTyrIleProProSerPheArg 95
Db 630 TTACCACTGACTTACCTCCTGAGCTTGTCTCCATCTTTGAGA 674

RESULT 9
US-09-764-875-91
; Sequence 91, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P202
; CURRENT APPLICATION NUMBER: US/09/764,875
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 955

```
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (550)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (934)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (942)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (944)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (948)
/ OTHER INFORMATION: n equals a,t,g, or c
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-91
```

```
Alignment Scores:
Pred. No.: 3.78e-38 Length: 955
Score: 334.00 Matches: 64
Percent Similarity: 81.05% Conservative: 13
Best Local Similarity: 67.37% Mismatches: 18
Query Match: 66.67% Indels: 0
DB: 11 Gaps: 0
```

US-09-720-934-2_COPY_215_310 (1-96) x US-09-764-875-91 (1-955)

```
QY 1 TrrAlvAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 390 TGGGCGCTTCTAGCCTACAGATTAAATATCGCAAAATTTAACTCTTGACAAA 449
QY 21 ThmetSerGlyHisLeuthrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 450 AGTATGAGTGATATCTTCAGGTTTCAAGTAGAAATGCCCTTCTTGCTCAAACTCTT 509
QY 41 ProGlnAlaGlnLeuAlaSerIleTpsanLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 510 TCTCAAACTGAGCTGCTACTATTGACTCTGCTGACGCTGATGATGACAGCTA 569
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 570 AAAGCAGAAAGATTATTCTTGCAATGACCTTACTGRCATGCGCAAAAGCTGACAGCCA 629
QY 81 LeuProValleuProProGlnIlyrIleProProSerPheArg 95
Db 630 TTACCACTGACTTACCTCCTGAGCTTGTTCTTCCTCCATCTTTGAGG 674
```

```
RESULT 10
US-09-764-881-54
/ Sequence 54, Application US/09764881
/ Publication No. US20020086821A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PT207
/ CURRENT APPLICATION NUMBER: US/09/764,881
/ PRIORITY FILING DATE: 2001-01-17
/ Prior application data removed - refer to PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 192
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 54
/ LENGTH: 955
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (550)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (934)
```

```
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (942)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (944)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (948)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-54
```

```
Alignment Scores:
Pred. No.: 3.78e-38 Length: 955
Score: 334.00 Matches: 64
Percent Similarity: 81.05% Conservative: 13
Best Local Similarity: 67.37% Mismatches: 18
Query Match: 66.67% Indels: 0
DB: 13 Gaps: 0
```

US-09-720-934-2_COPY_215_310 (1-96) x US-09-764-881-54 (1-955)

```
QY 1 TrrAlvAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 390 TGGGCGCTTCTAGCCTACAGATTAAATATCGCAAAATTTAACTCTTGACAAA 449
QY 21 ThmetSerGlyHisLeuthrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 450 AGTATGAGTGATATCTTCAGGTTTCAAGTAGAAATGCCCTTCTTGCTCAAACTCTT 509
QY 41 ProGlnAlaGlnLeuAlaSerIleTpsanLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 510 TCTCAAACTGAGCTGCTACTATTGACTCTGCTGACGCTGATGATGACAGCTA 569
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 570 AAAGCAGAAAGATTATTCTTGCAATGACCTTACTGRCATGCGCAAAAGCTGACAGCCA 629
QY 81 LeuProValleuProProGlnIlyrIleProProSerPheArg 95
Db 630 TTACCACTGACTTACCTCCTGAGCTTGTTCTTCCTCCATCTTTGAGG 674
```

```
RESULT 11
US-10-242-747-54
/ Sequence 54, Application US/10242747
/ Publication No. US2004005577A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PT207C1
/ CURRENT APPLICATION NUMBER: US/10/242,747
/ PRIORITY FILING DATE: 2002-09-13
/ PRIOR APPLICATION NUMBER: 09/764,881
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 60/179,065
/ PRIOR FILING DATE: 2000-01-31
/ PRIOR APPLICATION NUMBER: 60/180,628
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: 60/214,886
/ PRIOR FILING DATE: 2000-06-28
/ PRIOR APPLICATION NUMBER: 60/217,487
/ PRIOR FILING DATE: 2000-07-11
/ PRIOR APPLICATION NUMBER: 60/225,758
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/220,963
/ PRIOR FILING DATE: 2000-07-26
/ PRIOR APPLICATION NUMBER: 60/217,496
/ PRIOR FILING DATE: 2000-07-11
/ PRIOR APPLICATION NUMBER: 60/225,447
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/218,290
/ PRIOR FILING DATE: 2000-07-14
/ Remaining Prior Application data removed - See File Wrapper or PALM.
```


LENGTH: 3976
TYPE: DNA
ORGANISM: Aspergillus nidulans
US-10-369-493-36727

Alignment Scores:

Pred. No.:	1,37e-20	Length:	3976
Score:	218.50	Matches:	43
Percent Similarity:	62.89%	Conservative:	18
Best Local Similarity:	44.33%	Mismatches:	35
Query Match:	43.61%	Indels:	1
DB:	16	Gaps:	1

US-09-720-934-2_COPY_215_310 (1-96) x US-10-369-493-36727 (1-3976)

QY 1 TYPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAnsSerHisAspLys 20

DB 1043 TGGTTATTTCTCCGAGAGAACTCCAGTTTGACAACACTACTACTACTGTCGATACG 1102

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40

DB 1103 ACGAATCTGGGTTATACCGCGATCAAGCTGTACCGTTCTTACGAAGCGCGACGTA 1162

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheSerAspIleAspGlnAspGlyLysLeu 60

DB 1163 CCCAGAGAGACCTTGGCGCAGATTGGCATCTTGGCATTTGATGCTGATGGCCAGTGA 1222

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGln 79

DB 1223 ACAGGAGTAGTGTGGCGGTTGGCATGTATTGATTCGCCAGCAATTCACCAACAGGGC 1282

QY 80 ProLeuProValLeuProProGluTyrIleProProSerPheArgArg 96

DB 1283 CCCCTACCGCAGACTTACCAACCCGCTTGATACCCCAAGTAGTGGCGCA 1333

RESULT 14

US-09-884-441-60

Sequence 60, Application US/09884441

Patent No. US20020119158A1

GENERAL INFORMATION:

APPLICANT: Algate, Paul A.

APPLICANT: Carter, Darrick

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

FILE REFERENCE: 210121.462C7

CURRENT APPLICATION NUMBER: US/09/884,441

NUMBER OF SEQ ID NOS: 489

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 60

LENGTH: 480

TYPE: DNA

ORGANISM: Homo sapien

US-09-884-441-60

Alignment Scores:

Pred. No.:	2.41e-20	Length:	480
Score:	208.00	Matches:	44
Percent Similarity:	63.16%	Conservative:	16
Best Local Similarity:	46.32%	Mismatches:	33
Query Match:	41.52%	Indels:	2
DB:	9	Gaps:	2

US-09-720-934-2_COPY_215_310 (1-96) x US-09-884-441-60 (1-480)

QY 1 TYPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAnsSerHisAspLys 20

DB 56 TGGGCTATTACATCTGAAGACGTAAGTATGAATGAACAGTTGATACCTC---AAA 112

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40

DB 113 CCTTACGAGGATTACATTAACAGGTATCAAGCCCGTACTTTTCTTACAGTCAAGTCTG 172

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheSerAspIleAspGlnAspGlyLysLeu 60

DB 173 CCGGCCCCGGTTTACGTGAATAATGGCCCTTATCATGATCTGAACAGATGGCAAGATG 232

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

DB 233 GACCAAGCAAGATTCTTATAGCTATGAATCAAGTTAAAGTTGCAAGGCCCAACAG 292

QY 81 LeuProProValLeuProPro---GluTyrIleProProSerPhe 94

DB 293 CTGCCTGTAGTCTCCCTCTATCATGAACAACCCCTATGTTC 337

RESULT 15

US-09-907-969-60

Sequence 60, Application US/09907969

Publication No. US20030091580A1

GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer L.

APPLICANT: King, Gordon E.

APPLICANT: Algate, Paul A.

APPLICANT: Fling, Steven P.

APPLICANT: Retter, Marc W.

APPLICANT: Fanger, Gary Richard

APPLICANT: Reed, Steven G.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darrick

APPLICANT: Hill, Paul

APPLICANT: Albane, Earl

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

FILE REFERENCE: 210121.462C8

CURRENT APPLICATION NUMBER: US/09/907,969

NUMBER OF SEQ ID NOS: 596

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 60

LENGTH: 480

TYPE: DNA

ORGANISM: Homo sapiens

US-09-907-969-60

Alignment Scores:

Pred. No.:	2.41e-20	Length:	480
Score:	208.00	Matches:	44
Percent Similarity:	63.16%	Conservative:	16
Best Local Similarity:	46.32%	Mismatches:	33
Query Match:	41.52%	Indels:	2
DB:	10	Gaps:	2

US-09-720-934-2_COPY_215_310 (1-96) x US-09-907-969-60 (1-480)

QY 1 TYPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAnsSerHisAspLys 20

DB 56 TGGGCTATTACATCTGAAGACGTAAGTATGAATGAACAGTTTATTAACCTC---AAA 112

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40

DB 113 CCTTACGAGGATTACATTAACAGGTATCAAGCCCGTACTTTTCTTACAGTCAAGTCTG 172

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheSerAspIleAspGlnAspGlyLysLeu 60

DB 173 CCGGCCCCGGTTTACGTGAATAATGGCCCTTATCATGATCTGAACAGATGGGAAGATG 232

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

DB 233 GACCAAGCAAGATTCTTATAGCTATGAATCAAGTTAAAGTTGCAAGGCCCAACAG 292

QY 81 LeuProProValLeuProPro---GluTyrIleProProSerPhe 94

DB 293 CTGCCTGTAGTCTCCCTCTATCATGAACAACCCCTATGTTC 337

Search completed: July 1, 2004, 20:15:49

Fri Jul 2 10:23:38 2004

Job time : 300.52 secs

us-09-720-934-2_copy_215_310.rnpb

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 2887.87 Seconds
(without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 WAVPOSSRUKYRQLRNSHDK.....SGQPLPVLPPIPIPSFR 96

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEB=frame+ p2n.model -DEV=xlp
-Q/cgpn2_1/USPTO.spool/p/US09720334/funat 30062004 064540 13442/app.query.fasta_1.1386
-DB=ESR -QFMT=fastap -SUFPIX=est -MIMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAVISIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720334.@CGN 1.1 12431 @funat 30062004 064540 13442 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-FAPEXT=7 -Ygapop=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est4: *
13: gb_est3: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_hum: *
19: em_gss_hum: *
20: em_gss_hum: *
21: em_gss_hum: *
22: em_gss_hum: *
23: em_gss_hum: *
24: em_gss_hum: *
25: em_gss_hum: *
26: em_gss_hum: *
27: em_gss_hum: *
28: gb_gss1: *

29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	501	100.0	625	14	CF532126	CF532126 UI-M-GHO-
2	501	100.0	641	12	BM728997	BM728997 UI-E-BOI-
3	501	100.0	729	13	BU704308	BU704308 UI-M-FOO-
4	501	100.0	729	14	CF742748	CF742748 UI-M-HBO-
5	501	100.0	732	14	CA750495	CA750495 UI-M-FYO-
6	501	100.0	979	13	CF538122	CF538122 UI-M-GIO-
7	501	100.0	979	13	BO719508	BO719508 AGENCOURT
8	501	100.0	2079	11	BC013578	BC013578 Homo sapi
9	501	100.0	2126	11	BC020269	BC020269 Homo sapi
10	501	100.0	5385	11	BC062938	BC062938 Mus muscu
11	492	98.2	490	10	AM370310	AM370310 RCL-BT025
12	485	96.8	842	9	AL870708	AL870708 AL870708
13	484	96.6	776	13	BU119880	BU119880 603142511
14	483	96.4	776	13	BK758765	BK758765 BX758765
15	483	96.4	853	13	BK695683	BK695683 BX695683
16	474	94.6	646	12	BG161771	BG161771 de42408.X
17	471	94.0	819	14	CF741757	CF741757 UI-M-HBO-
18	462	92.2	445	10	BB849620	BB849620 BB849620
19	462	92.2	548	10	AM581906	AM581906 MR4-ST011
20	460	91.8	765	14	CF539131	CF539131 UI-M-GIO-
21	458	91.4	359	13	BY220693	BY220693 UI-M-GHO-
22	448	89.4	687	14	CB521525	CB521525 UI-M-GHO-
23	431	86.0	405	13	BY520359	BY520359 BY520359
24	425	84.8	907	13	BU128229	BU128229 603114019
25	421	84.0	585	14	CB518599	CB518599 UI-M-GHO-
26	418	83.4	486	12	BM152392	BM152392 TCBAPI888
27	417.5	83.3	1201	9	AL531589	AL531589 AL531589
28	417	83.2	606	10	AM910367	AM910367 UT80908.Y
29	405	82.8	722	14	CF533008	CF533008 UI-M-FYO-
30	407	81.2	911	13	BQ942708	BQ942708 AGENCOURT
31	376	75.0	724	10	BE373101	BE373101 601224527
32	352	70.3	886	13	BU317489	BU317489 603849519
33	349	69.7	425	9	AA154049	AA154049 m668405.x
34	349	69.7	639	10	BB242625	BB242625 BB242625
35	349	69.7	861	14	CD653097	CD653097 AGENCOURT
36	349	69.7	1550	11	AK042449	AK042449 Mus muscu
37	349	69.7	2352	11	AK034367	AK034367 Mus muscu
38	349	69.7	2641	11	AK029325	AK029325 Mus muscu
39	348	69.5	897	13	BO735187	BO735187 AGENCOURT
40	342	68.3	783	12	BG107389	BG107389 602290786
41	337	67.3	546	10	AW668653	AW668653 111193 MA
42	331	66.1	395	10	AW889826	AW889826 R55-NT003
43	331	66.1	574	10	BF816839	BF816839 MR2-CI012
44	316	63.1	627	9	AV896004	AV896004 AV896004
45	316	63.1	636	13	BM121860	BM121860 BM121860

ALIGNMENTS

RESULT 1
LOCUS CF532126 625 bp mRNA linear EST 12-SEP-2003
DEFINITION UI-M-GHO-cgu-d-05-0-UI.r1 NIH_BMAP_GHO Mus musculus cDNA clone
IMAGE:30356524 5', mRNA sequence.
ACCESSION CF532126
VERSION CF532126.1 GI:34584094
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 625)

QY 21 ThMeSerGIyhiSteuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 |||||
 Db 135 ACTATAGTGGACACTTAACGAGTCCCAAGCAAGAACTATCTTATGCAAGTCAAGTTTA 194
 QY 41 ProGlnAlaGlnLeuAlaSerIleTPanLeuSerAspIleAspGlnAspGlyLysIeu 60
 |||||
 Db 195 CCACAGGCTGAGCTGGCTTCAATATGGAATCTTTGACATTCATCAAGATGGAACCTT 254
 QY 61 ThrAlaGlnGluIupheIleuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 |||||
 Db 255 ACAGCAGAGAAATTATTCCTGGCAATGCACTCATTTAGTCTATGCTGGCCACCA 314
 QY 81 LeuProProValLeuProProGluTyrlleProProSerPheArgArg 96
 |||||
 Db 315 CTGCCACCTGTCTGCTGCTCCAGAAATACATCCACCTTTTGAAGA 362

RESULT 3

BU704308 729 bp mRNA linear EST 15-JUL-2003
 LOCUS UI-M-F00-bzs-j-21-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
 DEFINITION IMAGE:6406772 5', mRNA sequence.

ACCESSION BU704308
 VERSION BU704308
 KEYWORDS EST.
 SOURCE GI:23632308

ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 729)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.lnl.gov

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: PYX-5.

FEATURES
 source Location/Qualifiers

1..729
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6406772"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_id="NIH_BMAP_F00"
 /note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAAGAGGCC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 1.04e-46 Length: 729
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x BU704308 (1-729)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrlArgGlnLeuPheAsnSerHisAspLys 20
 |||||
 Db 259 TGGGCTGTGCTCTAGCATCAATCAAGCTGAATACAGCAGATTATTCACAGCCACGACAA 318
 QY 21 ThMeSerGIyhiSteuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 |||||
 Db 319 ACTATAGTGGACACTTAACAGGCTCCCAAGCAAGAACTATCTTCAATGCAATCAAGTTTA 378
 QY 41 ProGlnAlaGlnLeuAlaSerIleTPanLeuSerAspIleAspGlnAspGlyLysIeu 60
 |||||
 Db 379 CCCACAGCTCAGCTGCTTCATATGGAATCTTTGACATTCATCAAGATGGAACCTC 438
 QY 61 ThrAlaGlnGluIupheIleuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 |||||
 Db 439 ACTGCAAGAAATTATTCCTAGCTATGCACTCATTTAGTCTGCCATGTCTGCTACCCA 498
 QY 81 LeuProProValLeuProProGluTyrlleProProSerPheArgArg 96
 |||||
 Db 499 CTGCCACCTGTCTGCTGCTCCAGAAATACATCCCTCTTCAAGA 546

RESULT 4

CF742748 729 bp mRNA linear EST 10-OCT-2003
 LOCUS CF742748
 DEFINITION UI-M-HB0-clm-c-18-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
 IMAGE:30617825 5', mRNA sequence.

ACCESSION CF742748
 VERSION CF742748
 KEYWORDS EST.
 SOURCE GI:37639087

ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 729)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.owa.edu/distribution/mousefl.html

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 The following repetitive elements were found in this cDNA sequence: 589-705, >(CAG)n#Simple_repeat
 Seq primer: PYX-5.

FEATURES
 source Location/Qualifiers

1..729
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30617825"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_id="NIH_BMAP_HB0"
 /note="Organ: Eye; Vector: PYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES

source

Location/Qualifiers
 1..736
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30536000"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1lb="NIH BMAP G10"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: Bcor I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with Bcor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 1,05e-46 Length: 736
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x CFS38122 (1-736)

QY 1 TTPAlaValPProGlnSerSerArgLeuLYTYrArgGlnLeuPheAenSerHisAspLYs 20
 DB 185 TGGGCTGTGCTCAGTCATCAAGCGTGAATATACAGGAGTATATCAACGCCACGACAA 244
 QY 21 ThMeSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
 DB 245 ACTATAGTGGACACTTAACAGGTCGCCAGGCAAGAACTATCTCATGCAATCAAGTTTA 304
 QY 41 ProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLYsLeu 60
 DB 305 CCCAGGCTCAGCTGGCTTCAATATGGAATCTTTGCACATGATCAAGATGGAACAACTC 364
 QY 61 ThrAlaGlnGluIleuPheLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 DB 365 ACTGCAAGAGAAATTATCTCAGCTAGCTATGCACTAATGATGTCCTGCTGCTGACGCA 424
 QY 81 LeuProValLeuPProGlnuTYrIleProProSerPheArgArg 96
 DB 425 CTGCGCGCGCTCTGCTCCAGAAATACATCCCTCTCTTCACAGACA 472
 RESULT 7
 LOCUS BQ719508 979 bp mRNA linear EST 16-JUL-2002

DEFINITION

AGENCOURT 8219793 lupski sympathetic_trunk Homo sapiens cDNA clone
 IMAGE:6188111 5', mRNA sequence.

ACCESSION BQ719508
 VERSION BQ719508.1 GI:21858405
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>
 1 (bases 1 to 979)
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM13583 row: c column: 24
 High quality sequence stop: 643.

FEATURES

source

Location/Qualifiers
 1..979
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6188111"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /clone_1lb="Lupski sympathetic trunk"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1: NotI; Site 2: SalI. cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCG-3' and 5'-GACTAGTCTTAGATCGGAGCGGCGCCCTT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 1.58e-46 Length: 979
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x BQ719508 (1-979)

QY 1 TTPAlaValPProGlnSerSerArgLeuLYTYrArgGlnLeuPheAenSerHisAspLYs 20
 DB 353 TGGGCTGTGCTCAGTCATCAAGCTGAATATACAGGCAATTAATCAATAGTACACAA 412
 QY 21 ThMeSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
 DB 413 ACTATAGTGGACACTTAACAGGTGCCCAAGAACTATCTTATGCACTCAAGTTTA 472
 QY 41 ProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLYsLeu 60
 DB 473 CCACAGGCTCAGCTGGCTTCAATATGGAATCTTTGCACATGATCAAGATGGAACAACTT 532
 QY 61 ThrAlaGlnGluIleuPheLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 DB 533 ACAGCAGAGAAATTATCTGCGAATGACCTCATGATGATGATGCTGCGCAACCA 592

Qy 81 LeuProProValLeuProProGluTyrIleLeuProProSerPheArgArg 96
Db 593 CTGCACCTGCTCTGCTCCGAAATACATCCACCTCTTTTAAAGA 640

RESULT 8
LOCUS BC013578
DEFINITION Homo sapiens, similar to intersectin 1 (SH3 domain protein), clone IMAGE:3878242, mRNA.
ACCESSION BC013578
VERSION BC013578.1 GI:15488896
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2079)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.sbgc.stanford.edu>
Contact: (Dickson, Mark) mcddapx1.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES
source 1..2079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3878242"
/tissue_type="lung, large cell carcinoma"
/clone_id="NIH MGC_68"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN
Alignment Scores:
Pred. No.: 4.64e-46 Length: 2079
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) * BC013578 (1-2079)

Qy 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 774 TGGGCTGTCCTCAGTCATCAAGACTGAATACAGGCGCAATATTCATATGATCAGACAAA 833

Qy 21 ThMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIleu 40
Db 834 ACTATGATGACACCTTAACAGATCCCAAGCAAGAACTATTTATGACAGTCAAGTTTA 893

Qy 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60

Db 894 CCACAGGCTCAGCTGGCTTCATATGCAATCTTTCTACATTTGATCAAGATGAGAACTT 953

Qy 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 954 ACACGAGAGAAATTTATCTTGCAATGACACTTATGATGATGCTATGTCGGCAACCA 1013

Qy 81 LeuProProValLeuProProGluTyrIleLeuProProSerPheArgArg 96
Db 1014 CTGCACCTGCTCTGCTCCGAAATACATCCACCTCTTTTAAAGA 1061

RESULT 9
LOCUS BC020269
DEFINITION Homo sapiens, clone IMAGE:4899011, mRNA.
ACCESSION BC020269
VERSION BC020269.1 GI:17939664
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2126)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

FEATURES
source 1..2126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4899011"
/tissue_type="Pancicreas, epithelioid carcinoma"
/clone_id="NIH MGC_42"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN
Alignment Scores:
Pred. No.: 4.79e-46 Length: 2126
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAL Plate: 40 Row: n Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796
This clone has the following problem: no cloning site / microdeletion.

US-09-720-934-2_COPY_215_310 (1-96) x BC020269 (1-2126)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAsnSerHisAspLys 20

DB 824 TGGGCTGTTCTCCAGTCATCAACAGCAATGAAATACAGGCAATTATTCATATGATGACGAAA 883

QY 21 ThrMetSerGlyHisIleuThrGlyProGlnAlaArgThrIleuMetGlnSerSerLeu 40

DB 894 ACTATAGTGGACACTTAACAGGTCCCAAGCAAGCAACTATCTTATGACAGTCAAGTTTA 943

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGlnAspGlyLysLeu 60

DB 944 CCACAGGCTCAGCTGGCTTCAATATGCAATCTTTGACATTCATATCAAGTGAAGAACTT 1003

QY 61 ThrAlaGlnGlnPheIleuAlaMetHisIleuLysAspValAlaMetSerGlyGlnPro 80

DB 1004 ACAGCAGAGAGAAATTATCTCGCAATGACACCTCATGATGATGCTGTGGCCAAACA 1063

QY 81 LeuProProValLeuProProGlyTyrIleProProSerPheArgArg 96

DB 1064 CTGGCACTGCTCTGCTCCAGAAATACATTCACCTTTTGAAGA 1111

RESULT 10

BC062938 5385 bp mRNA linear HTC 11-DEC-2003

LOCUS Mus musculus interseitin (SH3 domain protein 1A), mRNA (cDNA clone IMAGE:6839463), containing frame-shift errors.

ACCESSION BC062938

VERSION BC062938.1 GI:3856052

KEYWORDS HTC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 5385)

AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schlier, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carinci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McKean, P.J., Mckernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalys, D.E., Scherz, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL 12477932

PubMed 2 (bases 1 to 5385)

REFERENCE Strausberg, R.

AUTHORS Direct Submission

TITLE Submitted (26-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>

Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu

Bonaldo, M.F., Akabou, I., Bail, T., Bail, J., Crouch, K., Davis, A., Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Smith, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: Plate: Row: Column: 0

This clone has the following problem: frame shifted.

FEATURES

source

1.5385

Location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6839463"

/issue_type="Brain"

/clone_lib="NIH_BMAP_GH0"

/lab_host="DH10B"

/note="Vector: pYX-ASC"

ORIGIN

Alignment Scores:

Pred. No.: 1.8e-45 Length: 5385

Score: 501.00 Matches: 96

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 11 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x BC062938 (1-5385)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAsnSerHisAspLys 20

DB 904 TGGGCTGTTCTCCAGTCATCAACAGCAATGAAATACAGGCAATTATTCATCAACCCAGCAAA 963

QY 21 ThrMetSerGlyHisIleuThrGlyProGlnAlaArgThrIleuMetGlnSerSerLeu 40

DB 964 ACTATAGTGGACACTTAACAGGTCCCAAGCAAGCAACTATTCATGACATCAAGTTTA 1023

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGlnAspGlyLysLeu 60

DB 1024 CCCAGGCTCAGCTGGCTTCAATATGCAATCTTTGACATTCATATCAAGTGAAGAACTC 1083

QY 61 ThrAlaGlnGlnPheIleuAlaMetHisIleuLysAspValAlaMetSerGlyGlnPro 80

DB 1084 ACTGAGAGAGAAATTATCTTACCTATGACACCTATGATGATGCCATGTGTGTCAGCCA 1143

QY 81 LeuProProValLeuProProGlyTyrIleProProSerPheArgArg 96

DB 1144 CTGGCGCCGCTCTGCTCCAGAAATACATCCCTCTTCTTCAAGA 1191

RESULT 11

AM370310/C

LOCUS AM370310 490 bp mRNA linear EST 04-FEB-2000

DEFINITION RCI-BT0255-041099-011-c09 BT0255 Homo sapiens cDNA, mRNA sequence.

ACCESSION AM370310

VERSION AM370310.1 GI:6874964

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 490)

AUTHORS HGCP <http://www.ludwig.org.br/ORESTES>.

TITLE The FAPESP/LICR Human Cancer Genome Project

JOURNAL Unpublished (1999)

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tbl=RC1&tbl2=RC1-BT0255-041099-011-c09&rc3=1999-10-04&rc4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 24
 Location/Qualifiers
 1. 490
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="B70255"
 /note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Alignment Scores:
 Pred. No.: 6,26e-46 Length: 490
 Score: 492.00 Matches: 95
 Percent Similarity: 98.96% Conservative: 0
 Best Local Similarity: 98.96% Mismatches: 1
 Query Match: 98.20% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AM370310 (1-490)

QY 1 TrrAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
 Db 337 TGGGCTGTTCTCTCAAGTCAATCAAGATCAAGGCAATTAATTCATGATGACACAA 278
 QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 Db 277 ACTATGATGGACACTTAACAGGTCGCCAAGCAGAACTATTTATGCACTCAAGTTTA 218
 QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
 Db 217 CCACAGGCTCAGCTGCGCTCAATATGGAATCTTTCTGACATTTGTTCAAGATGAAACTT 158
 QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 Db 157 ACACAGAGCAATTTATTCCTGGCAATGACCTCATTTGATGAGTATAGTCTGGGCCAACCC 98
 QY 81 LeuProBroValLeuProProGluTyrIleProProSerPheArgArg 96
 Db 97 CTGCCACCTGTCCTGCCCTCCAGAAATCACTTCCACTTTTACAGA 50

RESULT 12
 AL870708 842 bp mRNA linear EST 03-DEC-2003
 LOCUS AL870708 XGC-egg Silurana tropicalis cDNA clone TEG9141p04 5', mRNA
 DEFINITION sequence.
 ACCESSION AL870708
 VERSION AL870708.2 GI:38665439
 KEYWORDS EST.
 SOURCE Silurana tropicalis (western clawed frog)
 ORGANISM Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Silurana.
 1 (bases 1 to 842)
 Crounig, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
 Sanger Xenopus tropicalis EST project 2001 (11_2003)

JOURNAL
 COMMENT Unpublished (2003)
 On Sep 15, 2002 this sequence version replaced gi:22890973.
 Contact: Taylor R
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: TEG9141p04.plkBP6
 Sequencing primer: SP6
 This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.
 cDNA was oligo dt primed from 5ug of poly A+ RNA from egg.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI
 Host: Escherichia coli XL1-blue.
 Location/Qualifiers
 1. 842
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TEG9141p04"
 /dev_stage="egg"
 /lab_host="Escherichia coli XL1-blue"
 /clone_lib="XGC-egg"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from egg. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN
 Alignment Scores:
 Pred. No.: 8.5e-45 Length: 842
 Score: 485.00 Matches: 92
 Percent Similarity: 98.96% Conservative: 3
 Best Local Similarity: 95.83% Mismatches: 1
 Query Match: 96.81% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AL870708 (1-842)

QY 1 TrrAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
 Db 446 TGGGCTGATACACAGTCTTCAAGTTGAGTACAGACAGTTGTTCAACAGCCATGACAA 505
 QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 Db 506 ACCATGAGTGGGCAATTTAAACAGGCTTCAAGCAAGAACTATTTATGCACTCAAGTTTA 565
 QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
 Db 566 CCGAGAGCCAGCTAGTACGATGGAACCTTTCAAGACATTGACCTAGATGAGGAAAACTA 625
 QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 Db 626 ACGGCCAAGAATTTATTTAGTATGACATTATATGATGCGCCATGTCGGCAGCCA 685
 QY 81 LeuProBroValLeuProProGluTyrIleProProSerPheArgArg 96
 Db 686 CTTCCTCTCTGTTGCTCCTCGAGTATATTCCTCCCATCTTTTAAAGA 733

RESULT 13
 BU119880 786 bp mRNA linear EST 25-NOV-2002
 LOCUS BU119880
 DEFINITION 603142511F1 CSECHL16 Gallus gallus cDNA clone CHEST136f4 5', mRNA
 ACCESSION BU119880
 VERSION BU119880.1 GI:25329000
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

RESULT 15
BX695683/c
LOCUS BX695683 XGC-neurula Silurana tropicalis cDNA clone TNeu1ln07 3',
DEFINITION mRNA sequence.
ACCESSION BX695683
VERSION BX695683.1 GI:38357890
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Silurana.

REFERENCE 1 (bases 1 to 853)
AUTHORS Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL Unpublished (2003)
COMMENT Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK

FEATURES
source
1..853
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TNeu1ln07"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/clone_lib="XGC-neurula"
/note="Vector: PCS107; Site_1: EcoRI; Site_2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula. EcoRI-NotI cut cDNA was then ligated into PCS107 with EcoRI at the 5' end and NotI at the 3' end."
Host: Escherichia coli DH10B.
Vector: PCS107; Site_1: EcoRI; Site_2: NotI
Location/Qualifiers

ORIGIN

Alignment Scores:

Pred. No.: 1.46e-44 Length: 853
Score: 483.00 Matches: 92
Percent Similarity: 98.95% Conservative: 2
Best Local Similarity: 96.84% Mismatches: 1
Query Match: 96.41% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) * BX695683 (1-853)

QY 1 TrrAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 547 TGGGCTGTACCAAGTCTTCAAGTTGAGTACAGACAGTTGTCACAGCCACGACAA 488
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 487 ACCATGAGTGGGCACTTAACAGGTCCTCAAGCAGAACTATTCTTATGCACTCAAGTTTA 428
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
DB 427 CCGACAGCCAGCTAGCTACGATGGAACCTTCAGACATTGACCTAGATGGAAAACTA 368
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 367 ACGGCCAAGATTATTATTAGCTTATGACATTATATGATGGCCATGTCGCGACCA 308

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
DB 307 CTTCTCCAGTTCTGCTCTCTGAGTATATTCCTCCATCTTTTAGA 263

Search completed: July 1, 2004, 19:48:34
Job time : 2895.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 42.7143 Seconds
(without alignments)
792.522 Million cell updates/sec

Title: US-09-720-934-2_COPY_740_800
Perfect score: 333
Sequence: 1 VVVVYRRALYPPERSHDEI.....GGELKKGKTPPANYAEKIP 61

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cg2 1/USPTO.spool.p/US09720934/runat.30062004.064540.13455/app.query.fasta_1.1386
-DB=Issued Patents NA -QFMT=fastcap -SUFFIX=mi -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=JOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720934.@CGN 1.1.284 @runat.30062004.064540.13455 -NCPU=6 -ICPU=3
-NO MAMP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: Issued Patents NA.*
2: /cg2_6/prodata/2/ina/5A.COMB.seq.*
3: /cg2_6/prodata/2/ina/5B.COMB.seq.*
4: /cg2_6/prodata/2/ina/6A.COMB.seq.*
5: /cg2_6/prodata/2/ina/6B.COMB.seq.*
6: /cg2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215	64.6	2873	4	US-08-630-915A-193
2	140	42.0	970	4	US-09-833-381-1294
3	129	38.7	1392	1	US-08-475-894-3
4	129	38.7	1392	1	US-08-484-710-3
5	129	38.7	1392	2	US-08-484-709-3
6	129	38.7	1392	2	US-08-474-697-3
7	129	38.7	1659	1	US-08-475-894-1
8	129	38.7	1659	1	US-08-484-710-1
9	129	38.7	1659	2	US-08-484-709-1
10	129	38.7	1659	3	US-08-474-697-1
11	129	38.7	3143	4	US-08-671-354-1
12	128.5	38.6	2757	1	US-08-306-691B-48

13	128.5	38.6	2757	5	PCT-US93-06251-79	Sequence 79, Appl
14	124.5	37.4	2793	1	US-07-646-537B-1	Sequence 1, Appl
15	117	35.1	1277	4	US-08-630-915A-7	Sequence 7, Appl
16	116	34.8	691	4	US-08-630-915A-220	Sequence 220, Appl
17	116	34.8	901	5	PCT-US96-08950-1	Sequence 1, Appl
18	116	34.8	901	5	PCT-US96-09127-1	Sequence 1, Appl
19	116	34.8	1133	4	US-08-630-915A-35	Sequence 35, Appl
20	115.5	34.7	3111	4	US-10-014-882-1	Sequence 1, Appl
21	115.5	34.7	3518	4	US-10-014-882-3	Sequence 3, Appl
22	115	34.5	5325	4	US-10-164-595-17	Sequence 17, Appl
23	115	34.5	5334	4	US-10-164-595-21	Sequence 21, Appl
24	115	34.5	5581	4	US-10-164-595-19	Sequence 19, Appl
25	110	33.0	1661	2	US-08-815-176-2	Sequence 2, Appl
26	110	33.0	1661	4	US-09-197-344-2	Sequence 2, Appl
27	110	33.0	2100	3	US-08-938-830-2	Sequence 2, Appl
28	110	33.0	2100	3	US-09-020-222-2	Sequence 2, Appl
29	108	32.4	482	3	US-09-621-976-3867	Sequence 3867, Ap
30	108	32.4	3846	2	US-08-691-814B-3	Sequence 3, Appl
31	106	31.8	1199	4	US-08-630-915A-15	Sequence 15, Appl
32	106	31.8	1349	4	US-09-023-655-1176	Sequence 1176, Ap
33	106	31.8	1428	4	US-09-710-693-18	Sequence 18, Appl
34	106	31.8	1458	2	US-08-942-423-68	Sequence 68, Appl
35	106	31.8	1613	3	US-08-938-830-28	Sequence 28, Appl
36	106	31.8	1803	4	US-09-006-428A-18	Sequence 18, Appl
37	106	31.8	1858	4	US-09-006-428A-16	Sequence 16, Appl
38	106	31.8	2003	2	US-08-942-423-1	Sequence 1, Appl
39	106	31.8	2003	4	US-08-630-915A-25	Sequence 25, Appl
40	105	31.5	1382	4	US-09-820-005-1	Sequence 1, Appl
41	104.5	31.4	1240	4	US-09-833-381-888	Sequence 888, App
42	104	31.2	1636	4	US-08-630-915A-31	Sequence 31, Appl
43	104	31.2	1867	4	US-08-630-915A-13	Sequence 13, Appl
44	103.5	31.1	4762	4	US-09-300-958A-30	Sequence 30, Appl
45	102	30.6	4091	4	US-08-630-915A-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-08-630-915A-193
; Sequence 193, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOMLICKS, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-Apr-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090

```

; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIR
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2873 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; US-08-630-915A-193

Alignment Scores:
Pred. No.: 8.52e-21      Length: 2873
Score: 215.00           Matches: 36
Percent Similarity: 74.14%  Conservative: 7
Best Local Similarity: 62.07%  Mismatches: 15
Query Match: 64.56%      Indels: 0
DB: 4                   Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-630-915A-193 (1-2873)
QY 4 ValYrYrAgaIaLeuYrProPheGluSerHIsaSPgluIleThrIleGln 23
|||
Db 231 GTGAATTATAGAGCATTAATACCCCTTTGAACAGCAAGAACATGATGAGATGATTAAAT 290
QY 24 ProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeuGlyGly 43
|||
Db 291 TCTGAGATATATATTCACGTTGATGAAAAAACCGTAGAGAACCTGGTTGGCTTTATGCT 350
QY 44 GluLeuYrGlyYrPheProAlaAsnYrAlaGluYrIlePro 61
|||
Db 351 AGTTTCAAGAAATTTGGCTGGTTCCATGCAATTATGTAGAAAAAATGCCA 404

RESULT 2
US-09-833-381-1294/C
; Sequence 1294, Application US/0983381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. 6672186e1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1294
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1294

Alignment Scores:
Pred. No.: 1.01e-10      Length: 970
Score: 140.00           Matches: 26
Percent Similarity: 67.27%  Conservative: 11
Best Local Similarity: 47.27%  Mismatches: 16
Query Match: 42.04%      Indels: 2
DB: 4                   Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x US-09-833-381-1294 (1-970)
QY 7 ArgAlaLeuYrProPheGluSerArgSerHIsaSPgluIleThrIleGlnProGlyAsp 26
|||
Db 376 AAGGCTTGTATTTTTCAGCGAAGATATGATGAAATTTGGAAAAAGGGTAC 317
QY 27 IleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeuGlyGlyGluLeuYr 46
|||
Db 316 ATTGTGATTATACGAGAAAAA-----GAAGAAGATGGTGCTTGGATCTTGAAT 263
QY 47 GlyYrThrGlyTrpPheProAlaAsnYrAlaGluYrIlePro 61
|||
```

```

Db 262 GGGAAAAAGCCATTTTCTGCGCTTATGTGAGAGATTACT 218

RESULT 3
US-08-475-894-3
; Sequence 3, Application US/08475894
; Patent No. 5641748
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LANHVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,894
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-475-894-3

Alignment Scores:
Pred. No.: 5.98e-09      Length: 1392
Score: 129.00           Matches: 22
Percent Similarity: 64.91%  Conservative: 15
Best Local Similarity: 38.60%  Mismatches: 20
Query Match: 38.74%      Indels: 0
DB: 1                   Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-475-894-3 (1-1392)
QY 5 TyrTrYrAgaIaLeuYrProPheGluSerArgSerHIsaSPgluIleThrIleGlnPro 24
|||
Db 760 TACTGCAAGAAATATTTCCATATGAGGACAGCAATGATGATGACAATCAAGAA 819
QY 25 GlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeuGlyGlyGlu 44
|||
Db 820 GGAATATATGACTCTCATCATATAGACTGCATCGACGTAAGGCTGTGGAAAGAG 879
QY 45 LeuYrGlyYrThrGlyTrpPheProAlaAsnYrAlaGluYrIlePro 61
|||
Db 880 CTGAACGCAAGCAGAGCGGTGTTCCCGAATACCTGTGAAGTACTTCCA 930

RESULT 4
US-08-484-710-3
; Sequence 3, Application US/08484710
; Patent No. 5656438
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,709
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-192
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-484-709-3

Alignment Scores:
Alignment No.: 5,98e-09 Length: 1392
Score: 129.00 Matches: 22
Percent Similarity: 64.91% Conservative: 15
Best Local Similarity: 38.60% Mismatches: 20
Query Match: 38.74% Indels: 0
DB: 2 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) X US-08-484-709-3 (1-1392)
QY 5 TTTTATGATGAlaLeuTYrPheGluSerHisAspGluIleThrIleGlnPro 24
Db 760 TACTGGCAATGTAATTTCCATATGAGCACAGATGATGATGACATATAAGAA 819
QY 25 GLVAspIlaValMetValAspGluSerGlnThrGlyGluProGlyTYrPLeuGlyGlyGlu 44
Db 820 GGAGATATATGCTACTCTCATCATAGACCTGCATCCAGCCGTAAGCTGGTGGGAAGAGAG 879
QY 45 LeuLYsGlyLYsTYrThrGlyTYrPheProAlaAsnTYrAlaGluLYsIlePro 61
Db 880 CTGAAACGGACAGACGAGCGCTGTTCCCGCATTAACCTCGTAAGATTACTTCCA 930

RESULT 6
US-08-474-697-3
; Sequence 3, Application US/08474697
; Patent No. 6171800
; GENERAL INFORMATION:
; APPLICANT: Yen-ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,697
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941

```

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-474-697-3

Alignment Scores:
Pred. No.: 5.98e-09
Score: 129.00
Percent Similarity: 64.91%
Best Local Similarity: 38.60%
Query Match: 38.74%
DB: 3
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-474-697-3 (1-1392)

QY 5 TyTyTArGAlaLeuTyTPrPhGluSerArGserHisAspGluIleThrIleGlnPro 24
Db 760 TACTGCAAGTAATATTTCCATATGAGCAGACAGATGATGATGATGACATCAAAAGA 819

QY 25 GlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrLeuGlyGlyGlu 44
Db 820 GGAGATATAGTCACTCTCATATATAGAGACTGCATCGACGTAGGCTGTGGGAAAGAGAG 879

QY 45 LeuTyGlyTyrThrGlyTyrPhProAlaAsnTyraGluIlePro 61
Db 880 CTGAACGGCAGACAGAGGCGTGTCCCGATTAATTCTGTAAGTTACTTCCA 930

RESULT 7
US-08-475-894-1
Sequence 1, Application US/08475894
Patent No. 5641748
GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,894
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-475-894-1

Alignment Scores:
Pred. No.: 7.53e-09
Length: 1659

Score: 129.00
Percent Similarity: 64.91%
Best Local Similarity: 38.60%
Query Match: 38.74%
DB: 1
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-475-894-1 (1-1659)

QY 5 TyTyTArGAlaLeuTyTPrPhGluSerArGserHisAspGluIleThrIleGlnPro 24
Db 1027 TACTGCAAGTAATATTTCCATATGAGCAGACAGATGATGATGATGACATCAAAAGA 1086

QY 25 GlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrLeuGlyGlyGlu 44
Db 1087 GGAGATATAGTCACTCTCATATATAGAGACTGCATCGACGTAGGCTGTGGGAAAGAGAG 1146

QY 45 LeuTyGlyTyrThrGlyTyrPhProAlaAsnTyraGluIlePro 61
Db 1147 CTGAACGGCAGACAGAGGCGTGTCCCGATTAATTCTGTAAGTTACTTCCA 1197

RESULT 8
US-08-484-710-1
Sequence 1, Application US/08484710
Patent No. 5656438
GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,710
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-484-710-1

Alignment Scores:
Pred. No.: 7.53e-09
Score: 129.00
Percent Similarity: 64.91%
Best Local Similarity: 38.60%
Query Match: 38.74%
DB: 1
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-484-710-1 (1-1659)

QY 5 TyTyTArGAlaLeuTyTPrPhGluSerArGserHisAspGluIleThrIleGlnPro 24
Db 1027 TACTGCAAGTAATATTTCCATATGAGCAGACAGATGATGATGATGACATCAAAAGA 1086

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DIGILLO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 2757 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-79

Alignment Scores:
Pred. No.: 1.73e-08 Length: 2757
Score: 128.50 Matches: 22
Percent Similarity: 67.92% Conservative: 14
Best Local Similarity: 41.51% Mismatches: 16
Query Match: 38.59% Indels: 1
DB: 5 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x PCT-US93-06251-79 (1-2757)

QY 7 ArgAlaLeuTyPrpHeGluSerArgSerHisAspGluIleThrIleGlnProGlyAsp 26

Db 2328 AAAGCCGGTACACTTTCGCCCGTACCGCTTCAGAGCTGTCGCTCAAGAGGGGTAC 2387

QY 27 lIeValMetValAspGluSerGlnThrGlyGluProGlyTTrpLeuGlyGlyIuLeuLys 46

Db 2388 ATCATCAAGATC---CTTAACAAGAGGACAGCAAGCGTGGCGAGGGGAGATCTAT 2444

QY 47 GlyLysThrGlyTTrpHeProAlaAsnTyAlaGluLys 59

Db 2445 GGCCGGGTGGCTGGTTCCTGCCAAGTACGTGAAGAA 2483

RESULT 14
US-07-646-537B-1
Sequence 1, Application US/07646537B

Patent No. 5348664

GENERAL INFORMATION:
APPLICANT: Barbacid, Mariano

TITLE OF INVENTION: Yav Proto-Oncogene Protein

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company

STREET: P.O. Box 4000

CITY: Princeton

STATE: New Jersey

COUNTRY: U.S.A.
ZIP: 08543-4000

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/646,537B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gaul, Timothy J.

REGISTRATION NUMBER: 33,111
REFERENCE/DOCKET NUMBER: DC10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 921-5901
TELEFAX: (609) 921-4526
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 14..2545
US-07-646-537B-1

Alignment Scores:
Pred. No.: 6.55e-08 Length: 2793
Score: 124.50 Matches: 21
Percent Similarity: 67.92% Conservative: 15
Best Local Similarity: 39.62% Mismatches: 16
Query Match: 37.39% Indels: 1
DB: 1 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x US-07-646-537B-1 (1-2793)

QY 7 ArgAlaLeuTyPrpHeGluSerArgSerHisAspGluIleThrIleGlnProGlyAsp 26

Db 2372 AAAGCCGGTACACTTTCGCCCGGACAGAGTGGAGTGCCTTAAGAGGGGTGAT 2431

QY 27 lIeValMetValAspGluSerGlnThrGlyGluProGlyTTrpLeuGlyGlyIuLeuLys 46

Db 2432 ATCATCAAGATC---CTCAATAAGAGGACAGCAAGCGTGGCGTGGGAGATCTAC 2488

QY 47 GlyLysThrGlyTTrpHeProAlaAsnTyAlaGluLys 59

Db 2489 GGCCGGATCGGCTGGTTCCTTCAAGTACGTGAAGAA 2527

RESULT 15
US-08-630-915A-7
Sequence 7, Application US/08630915A

Patent No. 6309820

GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.

APPLICANT: HOFFMAN, No. 6309820h

APPLICANT: KAY, Brian K.

APPLICANT: MCCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL

TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND

NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penite & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8644/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1277 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-915A-7

Alignment Scores:
Pred. No.: 2.74e-07
Score: 117.00
Percent Similarity: 57.81%
Best Local Similarity: 43.75%
Query Match: 35.14%
DB: 4

Length: 1277
Matches: 28
Conservative: 9
Mismatches: 19
Indels: 8
Gaps: 3

US-09-720-934-2_COPY_740_800 (1-61) * US-08-630-915A-7 (1-1277)

QY	2	LySValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr	21
DB	155	CAAGTTAAAGCTTCACAGAGCTCTATATACATTGAACTCCAGACTCCAGATGAATTATAC	214
QY	22	IlleGlnProGlyAspIleValMetVal--AspGluSerGlnThrGlyGlnProGlyTyr	40
DB	215	TTTGAAAGAGGAGACATTATCTACATCATGACATGACATGATGATACC-----AGCTGG	265
QY	41	LeuGlyGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyr-----	56
DB	266	TGGAAAGGAGACATGCAAGGCGCAAGACAGACTGATCCGAGCAACTATGTGGCTGAGCGAG	325
QY	57	AlaGluLysIle	60
DB	326	GCAGATCCATT	337

Search completed: July 1, 2004, 19:53:46
Job time : 47.7143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 187.143 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_740_800
Perfect score: 333
Sequence: 1 VKVYVYRATYPRESRSHDEI.....GGELKGTGMFPANNAEKIP 61

Scoring table:
BLOSUM62 Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3163042 segs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frim+p2n.model -DEV=xip
-O/cgpn2_1/USPTO.spool.p/US09700934/runat.30062004.064541.13481/app.query.fasta_1.1386
-DB=Published Applications NA -QFMT=fastp -SUPFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62
-TRANS=human40.csl -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09700934 @CGN 1 1 1500 @runat.30062004.064541.13481
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORE=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YCAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: Published Applications NA:*
2: /cgpn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
3: /cgpn2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq:*
4: /cgpn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
5: /cgpn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgpn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq:*
7: /cgpn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgpn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgpn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgpn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgpn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgpn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgpn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgpn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgpn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq2:*
16: /cgpn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgpn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgpn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgpn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result *
No. Query Match Length DB ID Description

1	333	100.0	3319	11	US-09-764-875-88	Sequence 88, App1
2	333	100.0	3466	16	US-10-158-057-33	Sequence 33, App1
3	216	64.9	3746	11	US-09-764-875-176	Sequence 176, App
4	216	64.9	4210	9	US-09-764-868-125	Sequence 125, App
5	216	64.9	5828	13	US-10-398-885A-15	Sequence 15, App1
6	215	64.6	2873	9	US-09-879-957-193	Sequence 193, App
7	182	54.7	270	9	US-09-864-761-17127	Sequence 17127, A
8	182	54.7	287	9	US-09-864-761-26948	Sequence 26948, A
9	182	54.7	296	9	US-09-864-761-30453	Sequence 30453, A
10	182	54.7	475	9	US-09-864-761-311	Sequence 311, App
11	182	54.7	475	9	US-09-864-761-13884	Sequence 13884, A
12	182	54.7	480	9	US-09-864-761-10314	Sequence 10314, A
13	140	42.0	970	9	US-09-833-381-1294	Sequence 1294, App
14	140	42.0	2539	16	US-10-104-047-1655	Sequence 1655, App
15	139	41.7	581	10	US-09-918-995-14907	Sequence 6, App1
16	138	41.4	1558	9	US-09-939-825-6	Sequence 6, App1
17	138	41.4	1558	9	US-09-764-868-124	Sequence 124, App
18	138	41.4	1558	11	US-09-764-875-309	Sequence 309, App
19	129	38.7	386	13	US-10-085-783A-37196	Sequence 37196, A
20	129	38.7	386	16	US-10-242-535A-37196	Sequence 37196, A
21	129	38.7	400	13	US-10-085-783A-22639	Sequence 22639, A
22	129	38.7	400	16	US-10-242-535A-22639	Sequence 22639, A
23	129	38.7	1215	17	US-10-648-593-61	Sequence 61, App1
24	129	38.7	3143	14	US-10-144-621-1	Sequence 1, App1
25	129	38.7	3348	13	US-10-342-887-187	Sequence 187, App
26	129	38.7	3348	13	US-10-172-118-187	Sequence 187, App
27	128.5	38.6	518	9	US-09-796-692-5870	Sequence 5870, App
28	128.5	38.6	518	15	US-10-040-862-5870	Sequence 5870, App
29	128.5	38.6	518	16	US-10-057-475B-5870	Sequence 5870, App
30	128.5	38.6	518	16	US-10-154-884B-5870	Sequence 5870, App
31	128.5	38.6	2757	10	US-09-960-706-1059	Sequence 1059, App
32	124	37.2	397	9	US-09-924-035A-843	Sequence 843, App
33	124	37.2	397	13	US-09-770-423-865	Sequence 865, App
34	124	37.2	564	15	US-10-029-386-15	Sequence 15, App1
35	123	36.9	1336	17	US-10-437-963-53573	Sequence 53573, A
36	122	36.6	474	9	US-09-864-761-31955	Sequence 31955, A
37	122	36.6	474	9	US-09-864-761-15441	Sequence 15441, A
38	122	36.6	1890	13	US-10-424-599-12461	Sequence 12461, A
39	121	36.3	1592	17	US-10-437-963-60380	Sequence 60380, A
40	118	35.4	1702	13	US-10-424-599-19639	Sequence 19639, A
41	117.5	35.3	3027	13	US-10-087-192-1871	Sequence 1871, App
42	117.5	35.3	3092	13	US-10-042-865-15	Sequence 15, App1
43	117.5	35.3	3141	16	US-10-258-106-27	Sequence 27, App1
44	117.5	35.3	3538	13	US-10-362-892-36	Sequence 36, App1
45	117.5	35.3	3538	16	US-10-288-798-36	Sequence 36, App1

ALIGNMENTS

RESULT 1
US-09-764-875-88 Application US/09764875
Sequence 88, Application US/09764875
Publication No. US20040018969A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P202
CURRENT APPLICATION NUMBER: US/09/764,875
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 1249
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 88
LENGTH: 3319
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-875-88

Alignment Scores: 2.61e-39 Length: 3319
Pred. No.: 333.00 Matches: 61
Score: 33.30
Percent Similarity: 100.00% Conservative: 0


```

DB      605 TTGGTAATTATAGACATTATACCCCTTTGAAGCAAGGACCATGATGAGAGTTT 664
      |||
QY      23 GlnProGlyAspIleValMetValAspGluSerGlnThrGlyGlnProGlyTyrPleuGly 42
      |||
DB      665 AATTCTGGAGATATATATTCAGTTGATGAAAAAACGGTAGAGAACCTGGTGGCTTTAT 724
      |||
QY      43 GlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIlePro 61
      |||
DB      725 GGTAGTTTTCAGAGAAATTTGGCTGGTTCCATGCATTTATGTAGAAAAAATGCCA 781
      |||
RESULT 5
US-10-398-885A-15
; Sequence 15, Application US/10398885A
; Publication No. US20040053282A1
; GENERAL INFORMATION:
; APPLICANT: Sugita, Yuji
; APPLICANT: Hashida, Ryoichi
; APPLICANT: Ogawa, Kaoru
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Odayashi, Masaya
; APPLICANT: Saito, Hirohisa
; APPLICANT: Takahashi, Eiki
; TITLE OF INVENTION: Method of Testing For Allergic Diseases
; FILE REFERENCE: SHIMIZU-07907
; CURRENT APPLICATION NUMBER: US/10/398,885A
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: PCT/JP01/09937
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: JP 2000-314093
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 5828
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: CDS
; LOCATION: (7)..(5052)
; OTHER INFORMATION:
US-10-398-885A-15

Alignment Scores:
Pred. No.: 1,71e-21 Length: 5828
Score: 216.00 Matches: 36
Percent Similarity: 74.58% Conservative: 8
Best Local Similarity: 61.02% Mismatches: 15
Query Match: 64.86% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-10-398-885A-15 (1-5828)
QY      3 ValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIle 22
      |||
DB      2233 TTGGTGAATTATAGACATTATACCCCTTTGAAGCAAGGACCATGATGAGATGAGTTT 2292
      |||
QY      23 GlnProGlyAspIleValMetValAspGluSerGlnThrGlyGlnProGlyTyrPleuGly 42
      |||
DB      2293 AATTCTGGAGATATATATTCAGTTGATGAAAAAACGGTAGAGAACCTGGTGGCTTTAT 2352
      |||
QY      43 GlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIlePro 61
      |||
DB      2353 GGTAGTTTTCAGAGAAATTTGGCTGGTTCCATGCATTTATGTAGAAAAAATGCCA 2409
      |||
RESULT 6
US-09-879-957-193
; Sequence 193, Application US/09879957
; Patent No. US20020034755A1
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
```

```

; HOFMAN, No. US20020034755A1b
; KAY, Brian K.
; FOWKES, Dana M.
; MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/879,957
; FILING DATE: 13-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,915
; FILING DATE: 03-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2873 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 193:

US-09-879-957-193

Alignment Scores:
Pred. No.: 1e-21 Length: 2873
Score: 215.00 Matches: 36
Percent Similarity: 74.14% Conservative: 7
Best Local Similarity: 62.07% Mismatches: 15
Query Match: 64.56% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-09-879-957-193 (1-2873)
QY      4 ValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIleGln 23
      |||
DB      231 GTCGAATTATAGACATTATACCCCTTTGAAGCAAGGACCATGATGAGATGAGTTTAAAT 290
      |||
QY      24 ProGlyAspIleValMetValAspGluSerGlnThrGlyGlnProGlyTyrPleuGlyGly 43
      |||
DB      291 TCTGGAGATATATATTCAGTTGATGAAAAAACGGTAGAGAACCTGGTGGCTTTATAGGT 350
      |||
QY      44 GluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIlePro 61
      |||
DB      351 AGTTTTCAGAGAAATTTGGCTGGTTTCCATGCATTTATGTAGAAAAAATGCCA 404
      |||
RESULT 7
US-09-864-761-17127
; Sequence 17127, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
```



```
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 32
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.66
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: NT HIT: AF114488.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q15811, EVALUE 2.00e-45
; OTHER INFORMATION: EST_HUMAN HIT: AA773263.1, EVALUE 1.00e-112
US-09-864-761-26948

Alignment Scores:
Pred. No.: 4,93e-18 Length: 285
Score: 182.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) * US-09-864-761-26948 (1-286)
QY 30 ValAspGluSerGlnThrGlyGluProGlyTyrIleuGlyGlyLeuGlyGlyThr 49
DB 1 GTGATGAAAGCCAACTGGAGAACCCGGCTGGCTTGAGAGAGATTAAAGGAAAGACA 60
QY 50 GlyTyrPheProAlaSerTyrAlaGluTyrIlePro 61
DB 61 GGGTGGTTCCTCCGCAAACTATGACAGAAATCCCA 96

RESULT 9
US-09-864-761-30453
; Sequence 30453, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 30453
LENGTH: 297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000117.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.95
; OTHER INFORMATION: NT HIT: AF114487.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA773263.1, EVALUE 1.00e-112
; OTHER INFORMATION: SWISSPROT HIT: Q15811, EVALUE 2.00e-45
US-09-864-761-30453

Alignment Scores:
Pred. No.: 5.17e-18 Length: 297
Score: 182.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) * US-09-864-761-30453 (1-297)
QY 30 ValAspGluSerGlnThrGlyGluProGlyTyrIleuGlyGlyLeuGlyGlyThr 49
DB 1 GTGATGAAAGCCAACTGGAGAACCCGGCTGGCTTGAGAGAGATTAAAGGAAAGACA 60
QY 50 GlyTyrPheProAlaSerTyrAlaGluTyrIlePro 61
DB 61 GGGTGGTTCCTCCGCAAACTATGACAGAAATCCCA 96

RESULT 10
US-09-864-761-311
; Sequence 311, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```

```

PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 311
LENGTH: 475
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AF000193.1
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
US-09-864-761-311

Alignment Scores:
Pred. No.: 9,27e-18      Length: 475
Score: 182.00           Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65%      Indels: 0
DB: 9                    Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-09-864-761-311 (1-475)
QY 30 Valaspguserglnthrglyglnploglytrpleuglygyluleuylsglylysthr 49
Db 275 GTGATGAAAGCCAACTGGAAGAACCCGCTGCTTGAGAGAAATTAAAAAGAAAGACA 334
QY 50 G1YrPhpProAlaAsnYrAlaGluYsIlePro 61
Db 335 GGGTGTTCCCTGCACAACTATGACAGAGAAATCCCA 370

RESULT 11
US-09-864-761-13884
Sequence 13884, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aconica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761

```

```

CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 13884
LENGTH: 475
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000117.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.76
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.95
US-09-864-761-13884

Alignment Scores:
Pred. No.: 9,27e-18      Length: 475
Score: 182.00           Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65%      Indels: 0
DB: 9                    Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-09-864-761-13884 (1-475)
QY 30 Valaspguserglnthrglyglnploglytrpleuglygyluleuylsglylysthr 49
Db 275 GTGATGAAAGCCAACTGGAAGAACCCGCTGCTTGAGAGAAATTAAAAAGAAAGACA 334
QY 50 G1YrPhpProAlaAsnYrAlaGluYsIlePro 61
Db 335 GGGTGTTCCCTGCACAACTATGACAGAGAAATCCCA 370

```

```

RESULT 12
US-09-864-761-10314
; Sequence 10314, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10314
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000311.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.87
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 32
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.66
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
US-09-864-761-10314

Alignment Scores:
Pred. No.: 9,39e-18 Length: 480
Score: 182.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-09-864-761-10314 (1-480)

QY 30 ValAspGluSerGluThrGlyGluProGlyTyrPleuGlyGlyGluLeuYsglyLysThr 49
DB 243 GGGAGTGAAGGCGCAACCTGGAGAACCCGGCTGGCTGGAGAGATTAAAGGAAAGACA 302

QY 50 GlyTyrPheProAlaAsnTyrAlaGluYsilePro 61
DB 303 GGGTGGTTCCTGCAAACTATGCGAGAAATCCCA 338

RESULT 13
US-09-833-381-1294/c
; Sequence 1294, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1294
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1294

Alignment Scores:
Pred. No.: 4,36e-11 Length: 970
Score: 140.00 Matches: 26
Percent Similarity: 67.27% Conservative: 11
Best Local Similarity: 47.27% Mismatches: 16
Query Match: 42.04% Indels: 2
DB: 9 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x US-09-833-381-1294 (1-970)

QY 7 ArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIleGlnProGlyAsp 26
DB 376 AAGCGCTTGATTCCTTTTCAGCGCAAGCATGATGAGTTGAAATTTGGAAAGGGTGAC 317

QY 27 IleValMetValAspGluSerGlnThrGlyGluProGlyTyrPleuGlyGlyGluLeuYs 46
DB 316 ATTGTGATTATACACGAGAAAAA-----GAGGAGAGATGGTGGTTGATCTTTGAAT 263

QY 47 GlyLeThrGlyTyrPheProAlaAsnTyrAlaGluYsilePro 61
DB 262 GGGAAAAAAGCCATTTCCTCGCGCTTATGTGGAGAGATTACCT 218

RESULT 14
US-10-104-047-1655
; Sequence 1655, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: Helix Research Institute
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1655
; LENGTH: 2539
; TYPE: DNA

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 1835 Seconds
(without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_740_800
Perfect score: 333
Sequence: 1 VAVVYIRALYPRESHDEI.....GGLKCKTGWPANAEKIP 61

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xip
-O=/cgm2.1/USPTO.spool.p/US09703934/runat.30062004.064540.13442/app.query.fasta.1.1386
-DB=EST -OPMT=fastap -SUFFIX=xt -MINMATCH=0.1 -ICOPCU=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blowum62 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pico -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09703934 @CGM 1.1 12421 @runat.30062004.064540.13442 -NCPU=6 -ICPU=3
-NO MMAP -LARGEJOBBERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAVN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estda:*
2: em_esthm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est4:*
13: gb_est3:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt1:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	333	100.0	436	9	AL121153	AL121153 DKFZP762K
2	333	100.0	488	9	AA773823	AA773823 atf6d07.r
3	333	100.0	489	9	AA773263	AA773263 atf62e11.r
4	333	100.0	527	10	BB757493	BB757493 BB757493
5	333	100.0	573	14	CA874698	CA874698 K0934C04-
6	333	100.0	650	10	BB656585	BB656585 BB656585
7	333	100.0	657	10	AW950164	AW950164 EST362129
8	333	100.0	698	14	CF737796	CF737796 UI-M-HD0-
9	333	100.0	5385	11	BC062938	BC062938 Mus. muscu
10	329	98.8	608	14	CA511580	CA511580 UI-R-FU0-
11	320	96.1	703	14	CB527154	CB527154 UI-M-FY0-
12	319	95.8	673	14	CF735070	CF735070 UI-M-HB0-
13	317	95.2	443	9	AA492783	AA492783 v17307.r
14	302	90.7	600	14	CF738758	CF738758 UI-M-HD0-
15	301	90.4	551	9	AV669700	AV669700 AV669700
16	293	88.0	481	9	AA212646	AA212646 mus0c01.r
17	288	86.5	500	14	CD585892	CD585892 RK033A2H0
18	285	85.6	729	14	CF948685	CF948685 UI-M-HJ0-
19	281	84.4	688	14	CF729086	CF729086 UI-M-HD0-
20	277	83.2	456	9	AA313515	AA313515 EST185546
21	267	80.2	495	14	CA872596	CA872596 K0919C06-
22	244	73.3	148	9	AA670080	AA670080 atf79a03.r
23	239	71.8	337	14	CD586870	CD586870 RK038A3B0
24	237	71.2	201	14	CF906343	CF906343 A0446A03-
25	230	69.1	426	12	BC992086	BC992086 NR2-HT116
26	219	65.8	625	13	BU245991	BU245991 603592626
27	219	65.8	1133	13	BU144332	BU144332 603230249
28	217	65.2	833	12	BI872589	BI872589 UI-HF-BN0
29	216	64.9	653	14	CF136055	CF136055 UI-M-HD0-
30	216	64.9	700	13	BX481721	BX481721 DEFP686G
31	216	64.9	727	14	CF539609	CF539609 UI-M-GV0-
32	216	64.9	797	12	BC862028	BC862028 602795041
33	216	64.9	951	12	BM465259	BM465259 ACENOCURT
34	214	64.3	555	10	AA822841	AA822841 UG20F04.Y
35	201	60.4	774	13	BU316793	BU316793 603855209
36	198	59.5	384	9	AJ282752	AJ282752 4A3A-ABB-
37	198	59.5	681	12	BM596246	BM596246 170006874
38	183	55.0	699	13	BM216116	BM216116 BW216116
39	182	54.7	587	9	AV595297	AV595297 AV595297
40	179	53.8	144	14	X94499	X94499 RNSAP13M.Ra
41	178	53.5	527	14	CF729468	CF729468 UI-M-HD0-
42	176	52.9	1059	29	CNS057EX	AL132456 Tetradon
43	176	52.9	1072	29	CNS02AR2	AL18903 Tetradon
44	167	50.2	649	12	BJ103848	BJ103848 BJ103848
45	164	49.2	630	13	BQ179493	BQ179493 UI-M-EW0-

ALIGNMENTS

RESULT 1
LOCUS AL121153 436 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZP762K055.1 r1 762 (synonym: hme12) Homo sapiens cDNA clone
VERSION AL121153
ACCESSION AL121153
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 436)

AUTHORS Bloecker, H., Boecker, M., Brandt, P., Nemes, H.W., Gassenhuber, J. and Wiemann, S.
 TITLE EST (Bloecker, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: MIPS

MIPS
 Ingolstaedter Landstr. 1, D-85764 Neuberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de,
 sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the CDNA sequencing consortium of the German Genome Project.
 No 5' sequence available.
 This clone (DKFZp62K055) is available at the RZPD in Berlin.
 Please contact the RZPD, Ressourcenzentrum, Heidenweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 Location/Qualifiers

FEATURES

source 1..436
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp62K055"
 /tissue_type="melanoma (Mewo cell line)"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="762 (synonym: hmel2)"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Alignment Scores:

Pred. No.: 1.07e-31
 Score: 333.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%

Length: 436
 Matches: 61
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AL121153 (1-436)

QY 1 VallysVALVAllyTYTATGAlaleuTYrProphgIuSerArgSerHisAspGluIle 20
 Db 155 GTAAAGGAGGTGTATTACCGGCACTGTACCCCTTTGATCCAGAGCCATGATGAATC 214
 QY 21 ThrileGInProGlyAspIleValMetValAspGluSerGlnThnGlyGluProGlyTyr 40
 Db 215 ACTATCCAGCCAGGACATGATGCTGATGATGAAGCCAAACTGGAGAACCCGGCTGG 274
 QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
 Db 275 CTTGGAGAGAAATTAAGAGAAAGACAGGGTGGTTCCCTGCCTGCAAACTATGACAGAGAAATC 334
 QY 61 Pro 61
 Db 335 CCA 337

RESULT 2
 AA773823 488 bp mRNA linear EST 29-JAN-1998
 LOCUS AA773823
 DEFINITION af6d07.r1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1047949
 5' similar to SW:SP17_HUMAN Q15811 SH3 DOMAIN-CONTAINING PROTEIN SHP17. ;contains Alu repetitive element;; mRNA sequence.
 AA773823
 ACCESSION AA773823.1 GI:2825394
 VERSION
 KEYWORDS
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Matra, M., Martin, J., Moore, B., Schellenberg, K., Stepoe, M., Tan, F.,

TITLE Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 JOURNAL WashU-NCI human EST Project
 COMMENT Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available royalty-free through LML ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Putative full length read
 The vector to vector length is
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28m3 rev2 ET from Amersham
 High quality sequence stop: 479.
 Location/Qualifiers

FEATURES

source 1..488
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1047949"
 /tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
 /lab_host="DH10B"
 /clone_lib="Soares_NHMPu_S1"
 /note="Organ: mixed (see below); Vector: pTT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHW, pregnant uterus NbHPu, and fetal heart NbHH19w) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 48488-489479."

ORIGIN

Alignment Scores:

Pred. No.: 1.23e-31
 Score: 333.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%

Length: 488
 Matches: 61
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AA773823 (1-488)

QY 1 VallysVALVAllyTYTATGAlaleuTYrProphgIuSerArgSerHisAspGluIle 20
 Db 174 GTAAAGGAGGTGTATTACCGGCACTGTACCCCTTTGATCCAGAGCCATGATGAATC 233
 QY 21 ThrileGInProGlyAspIleValMetValAspGluSerGlnThnGlyGluProGlyTyr 40
 Db 234 ACTATCCAGCCAGGACATGATGCTGATGATGAAGCCAAACTGGAGAACCCGGCTGG 293
 QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
 Db 294 CTTGGAGAGAAATTAAGAGAAAGACAGGGTGGTTCCCTGCCTGCAAACTATGACAGAGAAATC 353
 QY 61 Pro 61
 Db 354 CCA 356

RESULT 3
 AA773263 489 bp mRNA linear EST 29-JAN-1998
 LOCUS AA773263
 DEFINITION af62e11.r1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1048556
 5' similar to SW:SP17_HUMAN Q15811 SH3 DOMAIN-CONTAINING PROTEIN SHP17. ;contains Alu repetitive element;; mRNA sequence.
 AA773263
 ACCESSION AA773263.1 GI:2824834
 VERSION

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 489)
Hillier, L., Allen, M., Bowler, L., Dubuque, T., Giesel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maita, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNM; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amer sham
High quality sequence stop: 445.

FEATURES
Location/Qualifiers
1..489
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1048556"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NhhMPU_S1"
/note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbhM, pregnant uterus NbhPU, and fetal heart NbhH19M) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1 M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

ORIGIN
Alignment Scores:
Pred. No.: 1,24e-31 Length: 489
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AA773263 (1-489)

QY 1 VallysyValValTYTYTARGAlaleuTYrProPhegiUserArgSerHisASPQluile 20
Db 175 GTAAAGGGGTGATTTACCGGCACTGATCCCTTGATCCAGAGGCATGATGAAATC 234
QY 21 ThrlleglnPrtogiyASPillevalMetValaspGluSerGlnThnglyGluPrGglATrp 40
Db 235 ACTATCCAGCCAGGAGACATGATGATGATGATGAAACCAACCTGAGGAACCCGGCTGG 294
QY 41 LeuglyGlyGluLeuLeuGlyTyrThrGlyTyrPheProAlaSerTYAlaGluysile 60
Db 295 CTTGAGGAGAAATTAAAGAAAGCAGGGTGGTTCCTCGCAACTATGACGAGAAATC 354
QY 61 Pro 61
Db 355 CCA 357

RESULT 4

BB757493
LOCUS 527 bp mRNA linear EST 16-OCT-2001
DEFINITION BB757493 RIKEN full-length enriched, melanocyte Mus musculus cDNA
clone G270085L11 3', mRNA sequence.
ACCESSION BB757493
VERSION BB757493
KEYWORDS BB757493.1 GI:16185846
SOURCE EST.
ORGANISM Mus musculus (house mouse)
REFERENCE Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 527)
AUTHORS Akimura, T., Aizawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Komno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Nutsaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Wachihi, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

TITLE Unpublished (2001)
JOURNAL Contact: Yoshihide Hayashizaki
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10): 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuhira, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiscapillary sequencer. Genome Res. 10 (11): 1757-1771 (2000)
Komno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2): 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

FEATURES
Location/Qualifiers
1..527
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G270085L11"
/cell_type="melanocyte"
/clone_lib="RIKEN full-length enriched, melanocyte"

ORIGIN
Alignment Scores:
Pred. No.: 1,36e-31 Length: 527
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x BB757493 (1-527)

Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Salto, T., Kiyosawa, H., Yamataka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y. Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

FEATURES

source

Location/Qualifiers
1..650
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D130043E15"
/tissue_type="spinal ganglion"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/lab_host="RIKEN full-length enriched, 12 days embryo spinal ganglion"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGCGCGCCGCACTGAGTTTCTTTTCTTTT 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGTTCTCGAGTTAATTAATTAATTCCTCCCTCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Alignment Scores:

Pred. No.: 1..78e-31 Length: 650
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x BB656585 (1-650)

QY 1 VallysvaVallyTYTtAgaAlaLeuTYrProPhelGluSerArgSerHisAspGluIle 20
DB 8 GTAAGTGTGATTAACCGAGCGCTGTACCCCTTGAATCCAGAACTCAGATGATC 67
QY 21 ThrlleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
DB 68 ACCATCCAGCCAGAGATATAGTCATGTGATGAAAGCCAGACTGAGAGCCAGATGG 127
QY 41 LeuGlyGlyGluLeuLeuGlyGlyTyrPheProAlaAsnTYrAlaGluIle 60
DB 128 CTGGAGAGAGCTGAAGGGAAGACGGATGTTCCCTCAACTATGAGAAAGATT 187
QY 61 Pro 61
DB 188 CCA 190
RESULT 7

AW950164 657 bp mRNA linear EST 01-JUN-2000
LOCUS
DEFINITION EST362129 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AW950164
VERSION AW950164.1 GI:8139701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 657)
Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaepard, R., Gay, C., Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and Quackenbush, J. Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray Unpublished (2000)
JOURNAL Contact: John Quackenbush
COMMENT The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 10
Seq primer: Reverse.

FEATURES

source

Location/Qualifiers
1..657
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone.lib="MAGE resequences, MAGE"
/note="Vector: pBluescriptSkm"

ORIGIN

Alignment Scores:

Pred. No.: 1..8e-31 Length: 657
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AW950164 (1-657)

QY 1 VallysvaVallyTYTtAgaAlaLeuTYrProPhelGluSerArgSerHisAspGluIle 20
DB 322 GTAAGTGTGATTAACCGAGCGCTGTACCCCTTGAATCCAGAACTGATGAATC 381
QY 21 ThrlleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
DB 382 ACTATCCAGCCAGAGACATAGTCATGTGATGAAAGCCAAACTGAGAAACCGGCTGG 441
QY 41 LeuGlyGlyGluLeuLeuGlyGlyTyrPheProAlaAsnTYrAlaGluIle 60
DB 442 CTGGAGAGAAATTAAAGGAAGACAGGGTGTTCCTCAACTATGAGAGAAATC 501
QY 61 Pro 61
DB 502 CCA 504
RESULT 8
CF737796 698 bp mRNA linear EST 10-OCT-2003
LOCUS
DEFINITION UT-M-HD0-ctc-t-11-0-UT-r1 NIH_BMAP_HD0 Mus musculus cDNA clone
IMAGE:30611818 5', mRNA sequence.
ACCESSION CF737796
VERSION CF737796.1 GI:37634133
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 698)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
Source Location/Qualifiers
 1. 698
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:36611818"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP HD0"
 /note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTATGAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,95e-31 Length: 698
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) * CFT37796 (1-698)

QY 1 VallystValVAllyTrrYrArGAlaLeuTyrProPhcGlUsErArGserHisAspGluIle 20
 Db 240 GTAAAGGGATATTAACCGAGGCGCTGTACCCCTTGATCCAGAAAGTCAGAGTGAAGTTC 299
 QY 21 ThrllleGlnProGlyAspIleValMeTValAspGluSerGlnThrGlyGluProGlyTrrP 40
 Db 300 ACCATCCAGCCAGAGATATGTCATGCTGATGAAGCCAGACTGAGAGCCAGATGG 359
 QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTrrPhcProAlaAsnTYrAlaGluLysIle 60
 Db 360 CTTGAGAGAGAGTGAAGGAGAGACGAGATGTTCCCTCGCAAACTATGACAGAAAGATT 419
 QY 61 Pro 61
 Db 420 CCA 422

RESULT 9
LOCUS BC062938
DEFINITION Mus musculus interectin (SH3 domain protein 1A), mRNA (cDNA clone IMAGE:6839463), containing frame-shift errors.
ACCESSION BC062938

VERSION BC062938.1 GI:38566052
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 5385)

AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.P., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Udén, T.B., Tschipynski, S., Carninci, P., Prange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullanly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Bulky, S.W., Vallalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Scherchenko, Y., Bonfield, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 5385)
 Strausberg, R.
 Direct Submission
 Submitted (26-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu, tom-casavant@uiowa.edu
 Bonaldi, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fischer, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Schaefer, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

ORIGIN
 Alignment Scores:
 Pred. No.: 2.63e-30 Length: 5385
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

FEATURES
source location/Qualifiers
 1. 5385
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6839463"
 /tissue_type="Brain"
 /clone_lib="NIH BMAP_GH0"
 /lab_host="DH10B"
 /note="Vector: pYX-Asc"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0
 This clone has the following problem: frame shifted.

Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x BC062938 (1-5385)
QY 1 ValysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
Db 2472 GTAAAGTGTATATTATACCGAGCGCTGTACCCCTTTGAATCCGAAAGTCACGATGAGATC 2531
QY 21 ThrTlGlnProGlyAspIleValMetValAspGluSerGlnThrGlyIleProGlyTyrP 40
Db 2532 ACCATCCAGCCAGAGATATAGTCATGTGATGAAAGCCAGACTGGAGCCAGATGG 2591
QY 41 LeuGlyGlyGluLeuLeuGlySerGlyThrGlyTyrPheProAlaAsnTyrAlaGlyIle 60
Db 2592 CTGGGAGGAGAGCTGAAGGAGAGACGGAGTGGTTCCTGCACAACTATGACGAAAAAGATT 2651
QY 61 Pro 61
Db 2652 CCA 2654
RESULT 10
CAS11580 608 bp mRNA linear EST 15-NOV-2002
LOCUS UI-R-FJO-cpy-m-13-0-UI.r1 UI-R-FJO Rattus norvegicus cDNA clone
DEFINITION UI-R-FJO-cpy-m-13-0-UI 5', mRNA sequence.
ACCESSION CAS11580
VERSION CAS11580.1 GI:25002534
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 608)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 REVERSE.
FEATURES
source
1..608
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-FJO-cpy-m-13-0-UI"
/tissue_type="embryo"
/dev_stage="embryo"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_1lb="UI-R-FJO"
/note="Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I;
UI-R-FJO is a cDNA library containing the following
tissue(s): rat embryo. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CATCTCTACT. This library
was created for the University of Iowa Program for Rat
Gene Discovery and Mapping (Val Sheffield, Bento Soares
and Tom Casavant")
ORIGIN
Alignment Scores:
Pred. No.: 5,17e-31 Length: 608
Score: 329.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.80% Indels: 0
DB: 14 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x CAS11580 (1-608)
QY 2 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 21
Db 135 AAAGTGTGTATTATACCGAGCGCTGTACCCCTTTGAGTCCAGAAATCATGACGATCACCC 194
QY 22 TleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyIleProGlyTyrPleu 41
Db 195 ATCCAGCCAGAGATATAGTCATGTGATGAAAGCCAGACCGAGACCGAGATGAGCTT 254
QY 42 GlyGlyGlyLeuLeuLeuGlySerGlyThrGlyTyrPheProAlaAsnTyrAlaGlyIlePro 61
Db 255 GGAAGGAACTGAAGGAGAGACGAGATGGTTCCTGCACAACTATGACGAAAGATTCCA 314
RESULT 11
CBS27154 703 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FVO-cfh-m-16-0-UI.r1 NIH BMAP_FVO Mus musculus cDNA clone
DEFINITION IMAGE:6849689 5', mRNA sequence.
ACCESSION CBS27154
VERSION CBS27154.1 GI:29360627
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 703)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
source
1..703
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6849689"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5gpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH BMAP_FVO"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.:	8.33e-30	Length:	703
Score:	320.00	Matches:	59
Percent Similarity:	96.72%	Conservative:	0
Best Local Similarity:	96.72%	Mismatches:	2
Query Match:	96.10%	Indels:	0
DB:	14	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x CBS27154 (1-703)

QY 1 VallyValValTYTYTARGAAlaLeuTYrProPheGluSerArgSerHisAspGluile 20
514 GTAAAGTGTATATTACCGAGCGCTGTACCCCTTTGATTCAGAAAGTCAGAGATGC 573
DB 21 ThrllleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYr 40
574 ACCATCCAGCCAGAGATATATGTCATGTGTGATGAAGCCAGACTGGAGACCCAGATGG 633
QY 41 LenglyGlyGluLeuLysGlyLysThryGlyTrpPheProAlaAsnTYrAlaGluLysile 60
634 CTGGAGAGAGCTGANAAGGAGACGAGATGTTCCCTGCANACTATGCAGAAAGATT 693
QY 61 Pro 61
DB 694 CCA 696

RESULT 12

CF735070

LOCUS CF735070 673 bp mRNA linear EST 10-OCT-2003
DEFINITION UT-M-HB0-ckc-d-09-0-UT.1 NIH BMAP HB0 Mus musculus cDNA clone
IMAGE:30615536 5', mRNA sequence.

ACCESSION CF735070

VERSION CF735070.1 GI:37631406

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE 1 (bases 1 to 673)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouse1.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1. .673
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"

FEATURES

source

/clone="IMAGE:30615536"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="NIH BMAP HB0"
/note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TTATGAGT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Alignment Scores:

Pred. No.:	1.05e-29	Length:	673
Score:	319.00	Matches:	59
Percent Similarity:	96.72%	Conservative:	0
Best Local Similarity:	96.72%	Mismatches:	2
Query Match:	95.80%	Indels:	0
DB:	14	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x CF735070 (1-673)

QY 1 VallyValValTYTYTARGAAlaLeuTYrProPheGluSerArgSerHisAspGluile 20
483 GTAAAGTGTATATTACCGAGCGCTGTACCCCTTTGATTCAGAAAGTCAGAGATGC 542
DB 21 ThrllleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYr 40
543 ACCATCCAGCCAGAGATATATGTCATGTGTGATGAAGCCAGACTGGAGACCCAGATGG 602
QY 41 LenglyGlyGluLeuLysGlyLysThryGlyTrpPheProAlaAsnTYrAlaGluLysile 60
603 CTGGAGAGAGCTGANAAGGAGACGAGATGTTCCCTGCANACTATGCAGAAAGATT 662
QY 61 Pro 61
DB 663 CCA 665

RESULT 13

AA492783

LOCUS AA492783 443 bp mRNA linear EST 25-JUN-1997
DEFINITION v177907.r1 Striatogene mouse testis (#937308) Mus musculus cDNA
clone IMAGE:918500 5', mRNA sequence.

ACCESSION AA492783

VERSION AA492783.1 GI:2222345

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE 1 (bases 1 to 443)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Stepec,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The Washu-HHMI Mouse EST Project

JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washu-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

COMMENT

Contact: Kiyoshi Naruse
Department of Biological Sciences
Graduate School of Science, University of Tokyo
Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: 81-3-5841-4443
Fax: 81-3-5841-4410
Email: naruse@biol.s.u-tokyo.ac.jp
This clone was isolated from OLN1 cell line cDNA library (OLb) 5'
end sequences.

FEATURES

Location/Qualifiers
1..551
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="HNI"
/db_xref="taxon:8090"
/clone="OLb10.079"
/clone_lib="OLN1 cell line cDNA library (OLb)"

ORIGIN

Alignment Scores:

Pred. No.:	1.46e-27	Length:	551
Score:	301.00	Matches:	51
Percent Similarity:	98.36%	Conservative:	9
Best Local Similarity:	83.61%	Mismatches:	1
Query Match:	90.39%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) * AV669700 (1-551)

QY	1	ValiysValiValiYrYrYrArgAlaLeuTyRProPheGluSerArgSerHisAspGluIle	20
Db	161	CTAAAGTGGTGTATACCGGGCGCTGTACCCCTTTGATGCCCGAGCACGATGAGATC	220
QY	21	ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr	40
Db	221	AGCATTTACTGTGAGACCTTATCATGGTGATGATCCAGACGGAGAGACCCGGCTGG	280
QY	41	LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle	60
Db	281	TTGGGCGGGAGCTCAGAGGTCCGACCGGCTTCCAGCTAATTATGCAGAACGATT	340
QY	61	Pro 61	
Db	341	CCG 343	

Search completed: July 1, 2004, 19:48:42
Job time : 1843 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 1345.81 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_908_966
Sequence: 1 VEGLOAQLYPMRAKKDNHL.....GEVQOGKGMPPKSYKILISG 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=/cgen2.1/USEPTO.spool.p/US09720934/runat.30062004.064539.13430/app.query.fasta_1.1386
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MIMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blcs62 -TRANS=human40.cdi -LIST=45
-LOCAL=GN-200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720934.@CEN.1.1.7509.@runat.30062004.064539.13430 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
GenEmbl.*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vtl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	328	100.0	3231	6 BD205037	BD205037 Isolated
2	328	100.0	3241	9 HSU61166	U61166 Human SH3 d
3	328	100.0	4103	5 AF032118	AF032118 Xenopus 1
4	328	100.0	4321	9 HSM86384	EX5381175 Homo sapi
5	328	100.0	5195	6 BD205035	BD205035 Isolated
6	328	100.0	5199	6 BD205033	BD205033 Isolated
7	328	100.0	5287	9 AF064243	AF064243 Homo sapi
8	328	100.0	5381	9 AF114488	AF114488 Homo sapi
9	328	100.0	5458	6 BD205034	BD205034 Isolated
10	328	100.0	6439	9 AF114487	AF114487 Homo sapi
11	328	100.0	7247	9 AF064244	AF064244 Homo sapi
12	328	100.0	52814	9 BS000183	BS000183 Pan trogl
13	328	100.0	100000	9 AP000193	AP000193 Homo sapi
14	328	100.0	114929	9 AP000050	AP000050 Homo sapi
15	328	100.0	125242	9 AF000311	AF000311 Homo sapi
16	328	100.0	142974	2 AC129075	AC129075 Felis cat
17	328	100.0	151516	9 AP000117	AP000117 Homo sapi
18	328	100.0	169774	2 AC138787	AC138787 Sus scrof
19	328	100.0	178801	2 AC121075	AC121075 Canis fam
20	328	100.0	184526	2 AC144367	AC144367 Papio anu
21	328	100.0	194157	2 AC144368	AC144368 Papio anu
22	328	100.0	203423	2 AC137544	AC137544 Gallus ga
23	328	100.0	224721	2 AC139632	AC139632 Bos tauri
24	328	100.0	235426	2 AC121074	AC121074 Canis fam
25	328	100.0	340000	9 AP001718	AP001718 Homo sapi
26	325	99.1	76179	2 AL606725	AL606725 Danio rer
27	325	99.1	104334	5 AL606751	AL606751 Zebrafish
28	325	99.1	163197	5 BX005416	BX005416 Zebrafish
29	325	99.1	182740	2 AC138436	AC138436 Danio rer
30	325	99.1	225535	2 BX470235	BX470235 Danio rer
31	323	98.5	3723	10 AF132478	AF132478 Mus muscu
32	323	98.5	3812	10 AF132672	AF132672 Rattus no
33	323	98.5	4025	10 AF127798	AF127798 Rattus no
34	323	98.5	5145	10 AF132481	AF132481 Mus muscu
35	323	98.5	175319	10 AC134837	AC134837 Mus muscu
36	323	98.5	181782	2 AC116970	AC116970 Rattus no
37	323	98.5	217131	10 AC126053	AC126053 Mus muscu
38	323	98.5	278501	2 AC098552	AC098552 Rattus no
39	323	98.5	279148	2 AC123507	AC123507 Rattus no
40	256	78.0	113146	2 AC138440	AC138440 Tetradon
41	256	78.0	127051	2 AC114895	AC114895 Tetradon
42	251	76.5	1391	9 AF001630	AF001630 Homo sapi
43	251	76.5	2004	9 AF038189	AF038189 Homo sapi
44	251	76.5	2873	6 AR175271	AR175271 Sequence
45	251	76.5	4447	6 AX428899	AX428899 Sequence

RESULT 1

ALIGNMENTS

BD205037 3231 bp DNA linear PAT 17-JUL-2003
LOCUS BD205037 Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205037
VERSION BD205037.1 GI:33014807
KEYWORDS JP 2002511267-A/5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3231)
AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 5 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL
COMMENT OS Homo sapiens (human)
PN JP 2002511267-A/5
PD 16-APR-2002
PR 16-APR-1998 JP 2000543610
PI JULIE R. KORENBERG, XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1..3231
FT 1..3231 /organism='Homo sapiens (human)'
ORIGIN
Location/Qualifiers
1..3231
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
Alignment Scores:
Pred. No.: 2.6e-37 Length: 3231
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x BD205037 (1-3231)
QY 1 ValGuglyLeuGlnAlaGlnAlaLeuTyrProTPrPaGAlaLyLyAspAsnHisLeu 20
DB 1214 GTGAGGGGGCTACCAAGCTCAAGCCTTATATCCTTGAGAGCCAAAGACACCACTTA 1273
QY 21 AsnPhaenLyAsnAspValIleThrValLeuGlnGlnAspMetTyrTrpPheGly 40
DB 1274 AATTTTAACAATAAATGATGATCATCAGCTCTGGAACGCAAGACATGTGGTGTGA 1333
QY 41 GluValGlnGlyGlnLySGlyTyrPheProLySerTyrValLyLeuIleSerGly 59
DB 1334 GAAGTTCAGAGTCAAGAGGTTGGTCCCAAGTCTTACGTGAACATCATTTCAAGG 1390
RESULT 2
LOCUS HSU61166 3241 bp mRNA linear PRI 23-JUL-1996
DEFINITION Human SH3 domain-containing protein SH3p17 mRNA, complete cds.
ACCESSION U61166
VERSION U61166.1 GI:1438932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3241)
AUTHORS Sparks,A.B., Hoffman,N.G., McConnell,S.J., Fowlkes,D.M. and
Kay,B.K.
TITLE Cloning of ligand targets: systematic isolation of SH3
domain-containing proteins
JOURNAL Nat. Biotechnol. 14 (6), 741-744 (1996)
MEDLINE 98294438
PUBMED 9630982
REFERENCE 2 (bases 1 to 3241)
AUTHORS Pirozzi,G., McConnell,S.J., Iveses,A. and Fowlkes,D.M.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1996) CYTOGEN Corp., 307 College Road East,
Princeton, NJ 08540, USA
FEATURES
Location/Qualifiers
source 1..3241
/organism='Homo sapiens'
/mol_type='mRNA'
/db_xref='taxon:9606'
/issue_type='bone marrow'
37..1599
/codon_start=1
/product='SH3 domain-containing protein SH3p17'
/protein_id='AAC50592.1'
/db_xref='GI:1438933'
/translation='MEAEFLKKEQERKIIIEKQKEBQRAQERDKQMEHVQOED
EHORPRKHEEKLRRESVKRQDGEKQEOAKLGLRPHQHPAPKPAVQAWST
AEKPLRTSAGNKKVYVYPALYPESHSHETITOPDLYMVDSQSGEGMIGEL
KGTGWFPNPAEKIPENEPVAPVYKVTDSAPAPKALAKETPAFLVTSSEPTTP
NNMADFSTWPTSTNEKRETNDWMAAQPSLTVPASQLEQKRSAPATATGSSPS
VLQGEKEVGLQAOALYPMRAKKNHNLNPNNDVITVLEQDMWFGVQOQKWFPK
SYVKLIISPIRSTSDSSSPASLKRVASPAKPVSGEETLAQVIASVYATGPEQ
LTLAAGQILIRKKNPGSGWEGELOAREKRGQICMPNANYKLLSPGSKITPEPPK
STLAAYQGVQMDYTAQNDELAFNKGQILINVLNKEDPPMWKGVNNGVGLPSNV
VKLTITMDPSQQ'
ORIGIN
Alignment Scores:
Pred. No.: 2.61e-37 Length: 3241
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x HSU61166 (1-3241)
QY 1 ValGuglyLeuGlnAlaGlnAlaLeuTyrProTPrPaGAlaLyLyAspAsnHisLeu 20
DB 886 GTGAGGGGGCTACCAAGCTCAAGCCTTATATCCTTGAGAGCCAAAGACACCACTTA 945
QY 21 AsnPhaenLyAsnAspValIleThrValLeuGlnGlnAspMetTyrTrpPheGly 40
DB 946 AATTTTAACAATAAATGATGATCATCAGCTCTGGAACGCAAGACATGTGGTGTGA 1005
QY 41 GluValGlnGlyGlnLySGlyTyrPheProLySerTyrValLyLeuIleSerGly 59
DB 1006 GAAGTTCAGAGTCAAGAGGTTGGTCCCAAGTCTTACGTGAACATCATTTCAAGG 1062
RESULT 3
LOCUS AF032118 4103 bp mRNA linear VRT 13-NOV-1998
DEFINITION Xenopus laevis intersectin mRNA, complete cds.
ACCESSION AF032118
VERSION AF032118.1 GI:2642624
KEYWORDS
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 4103)
AUTHORS Yamahata,M., Hoffman,N.G., Hardison,N.L., McPherson,P.S.,
Castagnoli,L., Cesareni,G. and Kay,B.K.

TITLE Intersectin, a novel adaptor protein with two Eps15 homology and five Src homology 3 domains
JOURNAL U. Biol. Chem. 273 (47), 31401-31407 (1998)
MEDLINE 99030416
PUBMED 9813051
2 (bases 1 to 4103)
AUTHORS Hoffman, N.G., Hardison, N.L., Yamabhai, M. and Kay, B.K.
TITLE Direct Submision (30-Oct-1997) Pharmacology, University of Wisconsin, 1300
JOURNAL Submitted (30-Oct-1997) University of Wisconsin, 1300
University Ave, Madison, WI 53706-1532, USA
Location/Qualifiers
1. 4103
FEATURES
source

CDs
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/cell_type="oocyte"
193. .4005
/function="involved in endocytosis"
/note="EH domain and SH3 domain containing protein;
similar to intersectin binding proteins Ibp1, encoded by
GenBank Accession Number AF057285, and Ibp2, encoded by
GenBank Accession Number AF057286, and mouse
Rab/Rip, encoded by GenBank Accession Number AF057287"
/codon_start=1
/product="intersectin"
/protein_id="AAC73068.1"
/db_xref="GI:2642625"
/translation="MAQFGTPFGGNDIMATVEERAKDQDFHGLKPTAGYITGDQA
RNFFLQSGLPQVLAQIWMALDMMNDMDQLEFSLAMLIKLGQYPLSLPSNM
LKOPVMPAAVAFVFGMSGIVGIPLAAPVPMPSIPVAGMSPLVSVPTVPLPSN
GAPAVIOHPAPVAFVHSAITLPRSSSGFSGSVAGSOINTLQKAGSPDPAPLVAVWAPS
SSRLKRLQFNSGDKTMSGNLTSPQAKRTILMOSLIFOSQATITWISDIDQKLTAE
EPLIAHLIDVANSQGLPPLIPPEIPSPFRVRSGLTMSVSVQRLPEEPER
EBQNDADKTLPTFEDKRENPENRMELEKROALBOQRKEQLAQERAEQERK
ERRPOQERRKOOLEKOLEKREPERKEIRERAEKRELEAKRELEQERAEQERK
RREILNORNEQEDIVLAKAKKTELEPELEAKDKHOLEGLQDRIOLITOREHI
ESTNKRRELRALIEITLQOOLESOQLIKMIPKOSLIDOLKQVQNSLHSDILTL
KRALTEKEIGROQLRQDLQBEKEKTRAKLOEDIVFNQLEKRELYNQOEFKQDPE
TEKIKQKELEKTSLELDKLEEDKRELEODKLMQDVQVEERYVFOOEKEREES
VOCEVAKRELEIEDEKPNKPFHOPPEPGKLGQIIPMNTEKAPLTIQGVKVVYAL
YPDASHEDEITTEPGDILIMVDESQGESEGLGELKGTGMFPANVARMPESEPS
TKPBAETTSAGPPIVHVASPVAAPAFNTSTNSNMNADSSFTWTNTDKVESDMDYT
MAOPSLTVPSAGHQRORSAFPATVTVGSSPSPIVGQKREGLQALVPMARKNDT
HINFNKNDVITVLEQDDMMFGEVQOKMFPSPYKILSGPRKSTSIDTSSEBPA
SLKRVSSPAFKPAIIGEYITSMITYSNEQDITFOQDILIVIKDGMWMTVIEK
TGVPSPNYVRPKDSEAGSGGKSLGKKRELAQVIAASYAATAPEQLTLAPGQLILIR
KNPGGMMWEGELQARGKKRQIGWFPANVYALSPGTNKSPTPEPPKTSLEPTCCVIG
MDYIKONDELAFSGKGVINLVNKEDPDMWKGELNGLFESNVYKLTDDPSPQO
FLGVNPAAGIATGDRPFLIPFRDGPDLPRAPAPPLASVVMIKERCFIARPCED
MNKYININ"
misc_feature
253. .498
/note="encodes EHA domain"
misc_feature
850. .1101
/note="encodes EHB domain"
misc_feature
2386. .2571
/note="encodes SH3A domain"
misc_feature
2881. .3057
/note="encodes SH3B domain"
misc_feature
3148. .3324
/note="encodes SH3C domain"
misc_feature
3364. .3558
/note="encodes SH3D domain"
misc_feature
3607. .3786
/note="encodes SH3E domain"
ORIGIN
Alignment Scores:
Pred. No.: 3.4e-37 length: 4103
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AF032118 (1-4103)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTirArgAlaAlaLysAspAsnHisLeu 20
Db 2881 GTAGAGGCGCTTACGACACAGCGCTGTATCTTGGAGAGCAAGAAGCAACATCTT 2940
QY 21 AsnPhaAsnLysAspAspValIleThrValLeuGluGlnGlnAspMetTyrTrpPheGly 40
Db 2941 AATTTTACAAATAATATGATTTCACAGGCTTCTTGAACAGAGATATGTGTGTGGG 3000
QY 41 GluValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuIleSerGly 59
Db 3001 GAAGTTCAAGGTCAAAAGCGCTGTGTTCCAAATCTTATGTAAGCTTATATCCGCT 3057
RESULT 4
HSM806384
LOCUS HSM806384 4321 bp mRNA linear Fri 17-JUN-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686J17173 (from clone DKFZp686J17173).
ACCESSION BX538175
VERSION BX538175.1 GI:31874620
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 4321)
Lauber, J., Bahr, A., Mewes, H.W., Weill, B., Amlid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
TITLE Direct Submision
JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberger, GERMANY
COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp686J17173) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heuberweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
1. 4321
FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="21q22.11"
/clone="DKFZp686J17173"
/tissue_type="human uterus endochel primary cell culture"
/clone_lib="686 (synonym: hicc3). Vector psport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
1. 4321
/gene="DKFZp686J17173"
<1. 1886
/gene="DKFZp686J17173"
/note="Intersectin long isoform, N-terminus truncated,
differentially spliced"
/codon_start=3
/product="hypothetical protein"
/db_xref="GI:31874621"
/translation="FEFFFLVYVHLKLAPVIOPLPAPHPAATLPKSSGFSRSG
PGSQNTKLOKQSFVAVSPVAEWAIVQSSRLKYLKQLEFNSDKTMSGHLTPQAK
IMQSGSLPOLASLWNLSIDIDQDKLTAEFLIAHLIDVANSQGLPPLVPEYIP
PSFRVRSGLTMSVSVQRLPEEPERKEIRERAEKRELEAKRELEQERAEQERK
ELEKRRQALBOQRKEQERLAQERAEQERKEERQERKEQERKEQERKEQERKE
QREERKREIRERAEKRELEQERAEQERKEERQERKEQERKEQERKEQERKEQER
FELEAKNDKHQLEBGLQDRIOLITOREHISTNKRRELRALIEITLHQQOLESQOM
LGELLPKQILNDQLKQVQNSLHSDILIVIKDGMWMTVIEKTS
KLOEIDIPNNQLEKREIRNKOOLQOKSMERLEKQERKEIRLEKQERAEQERK
QORDKOWLHVQOQDEHORPKRIHEEKREKRESYVKDQGEKQEAQDKLGRLEHQ
HPEPAKPAVQAPWSTAKGRLTISAQENKVVYALVFPESRSHDEITTEPGDIWV

polya_signal 4267..4272
/gene="DKFZp686j17173"
polya_site 4289
/gene="DKFZp686j17173"

ORIGIN

Alignment Scores:

Pred. No.: 3.61e-37 Length: 4321
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x HSM606384 (1-4321)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaAlaGlyAspAsnHisLeu 20
DB 2289 GTGAGGGGGCTACAGCTCAAGCCCTATATCTTGAGAGGCCAAGCAACACCACTTA 2348

QY 21 AsnHeaenlyAsaAspValIleThyValLeuGluGlnGlnAspMetTTPTriPheGly 40
DB 2349 AATTTTAACAATAATGATGATCAACCTCGAAGCAAGACATGTGTGTTTGA 2408

QY 41 GluValGlnGlyGlnGlyTyrPheProlySerTyrValIleLeuIleSerGly 59
DB 2409 GAAGTTCAAGGTCAAGAGGTTTGTTCCCAAGCTTTCGTGAACATCATTTCAAGG 2465

RESULT 5
BD205035

LOCUS BD205035 5195 bp DNA linear PAT 17-JUL-2003
DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.

ACCESSION BD205035
VERSION BD205035.1 GI:33014805
KEYWORDS JP 2002511267-A/3.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 5195)

AUTHORS Korenberg,J.R. and Chen,X.N.

TITLE Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof

JOURNAL Patent: JP 2002511267-A 3 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL

COMMENT

OS Homo sapiens (human)
PN JP 2002511267-A/3
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 6D/082007
PI JULIE R KORENBERG,XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

CC Leukemia
CC and utilization thereof.
FH Location/Qualifiers
FT source 1..5195
FT /organism="Homo sapiens (human)".

FEATURES

source Location/Qualifiers
1..5195
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4.44e-37 Length: 5195
Score: 328.00 Matches: 59

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BD205035 (1-5195)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaAlaGlyAspAsnHisLeu 20
DB 2360 GTGAGGGGGCTACAGCTCAAGCCCTATATCTTGAGAGGCCAAGCAACACCACTTA 3019

QY 21 AsnHeaenlyAsaAspValIleThyValLeuGluGlnGlnAspMetTTPTriPheGly 40
DB 3020 AATTTTAACAATAATGATGATCAACCTCGAAGCAAGACATGTGTGTTTGA 3079

QY 41 GluValGlnGlyGlnGlyTyrPheProlySerTyrValIleLeuIleSerGly 59
DB 3080 GAAGTTCAAGGTCAAGAGGTTTGTTCCCAAGCTTTCGTGAACATCATTTCAAGG 3136

RESULT 6
BD205033

LOCUS BD205033 5199 bp DNA linear PAT 17-JUL-2003
DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.

ACCESSION BD205033
VERSION BD205033.1 GI:33014803
KEYWORDS JP 2002511267-A/1.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 5199)

AUTHORS Korenberg,J.R. and Chen,X.N.

TITLE Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof

JOURNAL Patent: JP 2002511267-A 1 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL

COMMENT

OS Homo sapiens (human)
PN JP 2002511267-A/1
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 6D/082007
PI JULIE R KORENBERG,XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

CC Leukemia
CC and utilization thereof.
FH Location/Qualifiers
FT source 1..5199
FT /organism="Homo sapiens (human)".

FEATURES

source Location/Qualifiers
1..5199
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 4.45e-37 Length: 5199
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BD205033 (1-5199)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaAlaGlyAspAsnHisLeu 20
DB 2329 GTGAGGGGGCTACAGCTCAAGCCCTATATCTTGAGAGGCCAAGCAACACCACTTA 2988


```
/product="interfectin short isoform"
/prot_id="AA029953.1"
/db_xref="GI:4898825"
/translation="MAQFPFGSLDIWAIIVEERAKHDQOFHSIKPISGFTTGOA
RNFPSGLPQVLAQIMALAMNDMDQVPSIMAKLIKTKLQYQALPAPVM
KOOPVASSAPFAGMGIAMPPLTAVAPVMSIPIVMSPTLVSPVPAVPLAN
GAPVLOPAPAPAPATLPKSSFSFGSOLNTLQKQASFDVAVPVAWAVP
QSRKLYKROLFPHSHDKTMSGLHNGPQRTIMOSLPLQALATWNLSDIDQCKLTA
EETILAMHLIDVMSQDLPPLVLPPEYIPPSFRVRSQSGISTVISTSVQRLPEEPV
LEBOQOLERKLEVTEDKRENERFENGLERKRLLEQKREQERLAQLAEQE
RKRREOERERKQLLEKQLEKRELEERREERRELEIRERAKLEKROLEWE
RNRRELNRKQEDIVLKAKKLLELEALNOKKLEGLQDIDICRLTQPO
EISTKSRRLTAETLHLOOQESQOMGRILPEKQINDOLKOVQCSLRDSIV
TIRALEAKRLARQHLRDOLEVEKTRKSLCEIDIPNNOLKELRIHNNQOLQKKS
MEARLKQKQERKILTEKQKEBAQRQERBKQMLEHVOEDERHPRKLIBEKL
KRESEVKKQKEERKQEAQDKGLRLEPHQEPKAPVQAPWSTAERKPLTISAQENV
KVVYRALYPFESRSHDEITIOGDIWVAGVWDSQTEPGMLGELGKTKGFPA
NVAEKIPENVPAPVPTDSTAPAPKALRETPAVATSESPSTPNMADFSST
WPTSTKEPTDNDMDAAQPSLTPVSAQLORSATPATATGSSPSVIGQEKVE
GLQALYPMRARKDMLNPNKDVITVLEQOMWGFQOGKMPKRSVVKLISGP
IRKSTMSDSSSPASIKRVASPAKPVYSGEFTIAMTYTESSEGDLTFQGDVIL
VTKKDDWMTGVDAGVPSNVKRLKDSGSGTAKTSLGKKEPIAQLAVIATYAT
GPEOLTLAPQOLILIRKQNGWEGELQARGKROIGMPANVVKLISGTSKITPT
EPKSTALAAVCVITFWYDYTAQNDDELAINFKQIILVNLKEDPDMWKGENVQVGLF
PSNVVVKLTMDPSQ"
```

ORIGIN

Alignment Scores:

Pred. No.: 4.62e-37
Score: 328.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%

Length: 5381
Matches: 59
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AF114488 (1-5381)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTrpArgAlaIleAlaIleLeu 20

Db 3005 GTGAGGGGGGTACAGGTCAAGCCCTATATCTTGGAGAGCCAAAAGAACACCACTTA 3064

QY 21 AsnPhaSnlySasnaSpvAlleThrValLeuGluGlnGlnAspMetTTPTrpPheGly 40

Db 3065 AATTTTACAAAATGATGTCATACCGTCTGGAACAGCAAGCATGTGGTTTGA 3124

QY 41 GluValGlnGlyGlnTyGlyTTPPhaProIysSerTyValIleIleSerGly 59

Db 3125 GAAGTCAAGGTCAAGAGGTTGGTCCCAAGCTTACGAAATCATTTACGGG 3181

RESULT 9

BD205034

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```
PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
PC C1201/68,
PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
CC and utilization thereof.
PH key Location/Qualifiers
FT source 1..5458
FT Location/Qualifiers
1..5458
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
```

ORIGIN

Alignment Scores:

Pred. No.: 4.7e-37
Score: 328.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%

Length: 5458
Matches: 59
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BD205034 (1-5458)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTrpArgAlaIleAlaIleLeu 20

Db 3003 GTGAGGGGGGTACAGGTCAAGCCCTATATCTTGGAGAGCCAAAAGAACACCACTTA 3062

QY 21 AsnPhaSnlySasnaSpvAlleThrValLeuGluGlnGlnAspMetTTPTrpPheGly 40

Db 3063 AATTTTACAAAATGATGTCATACCGTCTGGAACAGCAAGCATGTGGTTTGA 3122

QY 41 GluValGlnGlyGlnTyGlyTTPPhaProIysSerTyValIleIleSerGly 59

Db 3123 GAAGTCAAGGTCAAGAGGTTGGTCCCAAGCTTACGAAATCATTTACGGG 3179

RESULT 10

AF114487

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..6439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1..6439

gene

misc_feature 4649..4819 /note="encodes PH domain"
misc_feature 4895..5143 /note="encodes C2 domain"
ORIGIN

Alignment Scores:
Pred. No.: 6.47e-37 Length: 7247
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AF064244 (1-7247)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTrrPargAlaLysLysAspAsnHisLeu 20
Db 2843 GTGAGAGGGGCTACAGAGCTCAAGCCCTATATCTTGAGAGCCAAAGAACACCACTTA 2902

QY 21 AsnPhaenLysAsnAspValIleThrValLeuGluGlnGlnAspMetTrrPheGly 40
Db 2903 AATTTTAACAAAAATGATGTCATCACCGTCTCGAACAAGACATGTGTGTTTGA 2962

QY 41 GluValGlnGlyGlnLysGlyTrrPheProLysSerTyrValLysLeuIleSerGly 59
Db 2963 GAAGTTCAAGGTCAAGAGGGTGGTTCCTCCCAAGCTTACGTGAACATCATTTCAAGG 3019

RESULT 12
BS000183 52814 bp DNA linear PRI 07-OCT-2003
LOCUS
DEFINITION Pan troglodytes chromosome 22 clone:PTB-086H16, map 22, partial
sequence, complete sequence.
ACCESSION BS000183 GI:37537450
VERSION
KEYWORDS HTG.
SOURCE
ORGANISM Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE 1
AUTHORS The Chimpanzee Chromosome 22 Sequencing Consortium.
TITLE DNA sequence of chimpanzee chromosome 22 and its evolutionary
implications
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 52814)
AUTHORS Wang, S., Cai, Z., Wang, B., Zheng, H., Zhang, Y., Zhang, X., Zhu, G.,
Lu, G., Fu, G. and Chen, Z.
TITLE Direct Submision
JOURNAL Submitted (26-MAY-2003) Shengyue Wang, Chinese National Human
Genome Center at Shanghai, Genomic Sequencing; No.250 Bibo Road,
Zhang Jiang Hi-TECH Park, Shanghai 201203, CHINA
(E-mail:wangsy@chgc.sh.cn, URL:http://www.chgc.sh.cn,
Tel:86-21-50801919, Fax:86-21-50801922)
COMMENT The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
*Chinese National Human Genome Center at Shanghai, Shanghai, China;
*GFP, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KIRB Genome Research
Center, Daejeon, Korea;
*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.

Center: Chinese National Human Genome Center at Shanghai Center
code: CHGC
Web site: http://chgc.sh.cn
Contact: wangsy@chgc.sh.cn

Project Information
Center project name: The Chimpanzee Chromosome 22 Sequencing Project
Center clone name: PTB-086H16

Summary Statistics

Sequencing vector: pUC18, 100% of reads
Chemistry: Dye-terminator Big Dye and ET; 100% of reads Assembly
program: Phrap; version 0.990329
Consensus quality: 52712 bases at least Q40
Consensus quality: 52814 bases at least Q30
Consensus quality: 52814 bases at least Q20
Quality coverage: 9.2x

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30);
an attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at one
plasmid
subclone or more than one M13 subclone;
and the assembly was confirmed by restriction digest.

Source information:
The PRB1 chimpanzee BAC library was prepared from DNA isolated from
cultured cells established from the blood of a single male
chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers
(http://www.gsc.riken.go.jp).
VECTOR: pKS145
Sequence Quality Assessment:
This entry has been annotated with sequence
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in
10,000 bp.

Neighboring clones: PTB-301D17(left) and RP43-082007(right).

FEATURES
source
Location/Qualifiers
1..52814
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="PTB-086H16"
/clone_id="PRB1 chimpanzee BAC"

ORIGIN

Alignment Scores:
Pred. No.: 6.12e-36 Length: 52814
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BS000183 (1-52814)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTrrPargAlaLysLysAspAsnHisLeu 20
Db 43289 GTGAGAGGGGCTACAGAGCTCAAGCCCTATATCTTGAGAGCCAAAGAACACCACTTA 43348

QY 21 AsnPhaenLysAsnAspValIleThrValLeuGluGlnGlnAspMetTrrPheGly 40
Db 43349 AATTTTAACAAAAATGATGTCATCACCGTCTCGAACAAGACATGTGTGTTTGA 43408

QY 41 GluValGlnGlyGlnLysGlyTrrPheProLysSerTyrValLysLeuIleSerGly 59
Db 43409 GAAGTTCAAGGTCAAGAGGGTGGTTCCTCCCAAGCTTACGTGAACATCATTTCAAGG 43465

RESULT 13
AP000193 100000 bp DNA linear PRI 08-JAN-2000
LOCUS
DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,
clone 076C10-f32E9, segment 20/21, complete sequence.
ACCESSION AP000193 GI:4826584
VERSION AP000193.1
KEYWORDS HTG.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 100000)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Toroki,Y. and Sakaki,Y.
 TITLE Homo sapiens 2,083,744bp genomic DNA of 21q22.1 (REGION: D21S26-AML CLONE RANGE: Q78C10-f32B9)
 JOURNAL Published Only in Database (1999)
 REFERENCE 2 (bases 1 to 100000)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Toroki,Y. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAY-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)
 COMMENT E. coli transposon insertion: The present data does not contain E. coli transposon sequences which integrated in the original/previous sequences. We determined the boundary between the insertion and genomic sequences experimentally, removed the insertion sequences, reconstituted the present data. The sequencing project is supported by Japan Science Technology Corporation (JST) and The Institute of Physical and Chemical Research (RIKEN).
 FEATURES
 source
 1..100000
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.1"
 Alignment Scores:
 Pred. No.: 1,266-35 Length: 100000
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AP000193 (1-100000)
 QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTTPARGAlaLysLysAspAsnHisLeu 20
 DB 46054 GTGAGGGGCTCAAGCTCAAGCCCTATATCTTGGAGAGCCAAAAGACAACACTTA 46113
 QY 21 AenPheAsnLysAspAspValIleThrValLeuGluGlnGlnAspMetTrrPrrPheGly 40
 DB 46114 AATTTTAAACAAAATGATGTCTATCACCCTCGGAACAGACACTGTGTGGTTTGA 46173
 QY 41 GiuValGlnGlyGlnLysGlyTrrPrrPhePrrLysSerTYrValLysLeuLysSerGly 59
 DB 46174 GAAAGTTCAAGCTCAGAGGGTGTTCCTCCCAAGTCTTACGTGAACACTCATTTCAAGG 46230
 RESULT 14
 LOCUS AP000050 114929 bp DNA linear PRI 20-NOV-1999
 DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, segment 21/28, complete sequence.
 ACCESSION AP000050
 VERSION AP000050.1 GI:3132360
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 114929)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q

JOURNAL Published Only in Database (1998)
 REFERENCE 2 (bases 1 to 114929)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-1998) Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@gsc.riken.go.jp, Tel:0427-78-9732, Fax:0427-78-9561)
 COMMENT This sequence is conducted by Kitasato University JST sequencing laboratory as a JST sequencing team.
 Principal Investigator: Yoshiyuki Sakaki Ph.D.
 phone: +81-3-5449-5622, Fax : +81-3-5449-5445, sakaki@gsc.riken.go.jp
 Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D. The sequence is submitted by Human Genome Sequencing in ALIS project of JST
 Japan Science and Technology Corporation (JST)
 5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0078 Japan
 For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www-alis.tokyo.jst.go.jp/HGS/top.html>) or send email to webmaster@www-alis.tokyo.jst.go.jp.
 FEATURES
 source
 1..114929
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.1"
 /clone="130N6-149C3f_2"
 Alignment Scores:
 Pred. No.: 1,486-35 Length: 114929
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AP000050 (1-114929)
 QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTTPARGAlaLysLysAspAsnHisLeu 20
 DB 9467 GTGAGGGGCTCAAGCTCAAGCCCTATATCTTGGAGAGCCAAAAGACAACACTTA 9526
 QY 21 AenPheAsnLysAspAspValIleThrValLeuGluGlnGlnAspMetTrrPrrPheGly 40
 DB 9527 AATTTTAAACAAAATGATGTCTATCACCCTCGGAACAGACACTGTGTGGTTTGA 9586
 QY 41 GiuValGlnGlyGlnLysGlyTrrPrrPhePrrLysSerTYrValLysLeuLysSerGly 59
 DB 9587 GAAAGTTCAAGCTCAGAGGGTGTTCCTCCCAAGTCTTACGTGAACACTCATTTCAAGG 9643
 RESULT 15
 LOCUS AP000311 125242 bp DNA linear PRI 20-NOV-1999
 DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, D21S26-AML region, clone:130N6, complete sequence.
 ACCESSION AP000311
 VERSION AP000311.1 GI:4835680
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 125242)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Toroki,Y. and Sakaki,Y.
 TITLE Homo sapiens 125,242bp genomic DNA of 21q22.1
 JOURNAL Published Only in Database (1999)
 REFERENCE 2 (bases 1 to 125242)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

TITLE
JOURNAL

Fujiyama, A., Yada, T., Tohoki, Y. and Sakaki, Y.
Direct Submission
Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555,
Japan (E-mail: hattori@gsc.riken.go.jp,
URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,
Fax: 81-42-778-9924)

COMMENT
The sequence is a part of the data (ACCESSION No. AP000174 -
AP000194).

The sequencing project is supported by Japan Science Technology
Corporation (JST) and The Institute of Physical and Chemical
Research (RIKEN).

FEATURES
Location/Qualifiers
source
1. .125242
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="130N6"

ORIGIN

Alignment Scores:

Pred. No.:	1.63e-35	Length:	125242
Score:	328.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_908_966 (1-59) * AP000311 (1-125242)

QY	1	ValGlnGlyLeuGlnAlaGlnAlaLeuTyRProTPrArgAlaLysLysAspAsnHisLeu	20
DB	109467	GTGGAGGGGCTACGAGCTCAAGCCCTATATCTTGAGAGCCAAAGACCAACCACTTA	109526
QY	21	AsnPhenylLysAsnAspValIleThrValLeuGlnGlnAspMetTPrPheGly	40
DB	109527	AATTTTAAACAAAAATGATGATACCCGTCGAGACGACAGCATGTGTGTGGA	109586
QY	41	GluValGlnGlyGlnLysGlyTPrPheProLysSerTyRValLysLeuLeSerGly	59
DB	109587	GAACTTCAAGGTCGAAAGGTTGGTCCCAAGCTTACGTGAAGCACTATTTCAGGG	109643

Search completed: July 1, 2004, 16:15:13
Job time : 1396.81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 41.3138 Seconds
(without alignments)
792.522 Million cell updates/sec

Title: US-09-720-934-2_COPY_908_966
Perfect score: 328
Sequence: 1 VEGGLQALYPRAKKDNHL.....GEVQCKGKFPKSYKYLISG 59

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-O/cgn2_1/USPTO.spool.p/US09720934/unat.30062004.064540.13455/app.query.fasta_1.1386
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPOPT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934.@CEN 1 1 284 @unat.30062004.064540.13455 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=0 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	251	76.5	2873	4	US-08-630-915A-193
2	129.5	39.5	970	4	US-09-833-811-1294
3	125.5	38.3	2955	4	US-09-620-312D-676
4	125.5	38.3	3039	4	US-09-620-312D-675
5	124.5	38.0	5325	4	US-10-164-595-17
6	124.5	38.0	5334	4	US-10-164-595-21
7	124.5	38.0	5581	4	US-10-164-595-19
8	120.5	36.7	1458	2	US-08-942-423-68
9	120.5	36.7	2003	2	US-08-942-423-1
10	120.5	36.7	2003	4	US-08-630-915A-25
11	119.5	36.4	1968	4	US-09-023-655-1296
12	117.5	35.8	1867	4	US-08-630-915A-13

13	116.5	35.5	747	4	US-08-630-915A-39	Sequence 39, Appl
14	113.5	34.6	1392	1	US-08-475-894-3	Sequence 3, Appl
15	113.5	34.6	1392	1	US-08-484-710-3	Sequence 3, Appl
16	113.5	34.6	1392	2	US-08-484-709-3	Sequence 3, Appl
17	113.5	34.6	1392	2	US-08-474-697-3	Sequence 3, Appl
18	113.5	34.6	1659	1	US-08-475-894-1	Sequence 1, Appl
19	113.5	34.6	1659	1	US-08-484-710-1	Sequence 1, Appl
20	113.5	34.6	1659	2	US-08-484-709-1	Sequence 1, Appl
21	113.5	34.6	1659	2	US-08-474-697-1	Sequence 1, Appl
22	113.5	34.6	1659	3	US-08-671-354-1	Sequence 1, Appl
23	113	34.5	874	4	US-09-023-655-931	Sequence 931, App
24	111.5	34.0	1640	4	US-09-023-655-650	Sequence 650, App
25	111	33.8	2757	5	US-08-306-691B-48	Sequence 48, Appl
26	111	33.8	2757	5	PCT-US93-06251-79	Sequence 79, Appl
27	111	33.8	2793	1	US-07-646-537B-1	Sequence 1, Appl
28	108	32.9	1414	4	US-09-023-655-1356	Sequence 1356, Ap
29	108	32.9	5330	4	US-09-023-905A-1	Sequence 1, Appl
30	107.5	32.8	2074	4	US-08-630-915A-19	Sequence 19, Appl
31	107.5	32.8	2445	4	US-09-566-921-20	Sequence 20, Appl
32	107.5	32.8	4695	4	US-09-620-312D-379	Sequence 379, App
33	107	32.6	3456	4	US-09-023-905A-5	Sequence 5, Appl
34	107	32.6	3870	1	US-08-138-641-1	Sequence 1, Appl
35	107	32.6	3870	1	US-08-138-641-3	Sequence 1, Appl
36	107	32.6	3893	1	US-08-138-641-3	Sequence 1, Appl
37	107	32.6	3893	1	US-08-138-133-3	Sequence 3, Appl
38	107	32.6	4382	4	US-09-023-905A-3	Sequence 3, Appl
39	106.5	32.5	813	4	US-09-023-655-1238	Sequence 1238, Ap
40	106.5	32.5	933	1	US-09-641-640-1	Sequence 1, Appl
41	106.5	32.5	933	2	US-08-612-857-1	Sequence 1, Appl
42	106.5	32.5	984	4	US-09-765-298A-5	Sequence 5, Appl
43	106.5	32.5	1072	1	US-07-906-349A-2	Sequence 2, Appl
44	106.5	32.5	1072	1	US-08-167-035-5	Sequence 5, Appl
45	106.5	32.5	1072	1	US-08-167-035-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-08-630-915A-193
Sequence 193, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 48,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

```
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:
LENGTH: 2873 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-630-915A-193

Alignment Scores:
Pred. No.: 2,68e-28
Score: 251.00
Percent Similarity: 86.44%
Best Local Similarity: 74.58%
Query Match: 76.52%
DB: 4
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x US-08-630-915A-193 (1-2873)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTPrpArgAlaLysLysAspAsnHisLeu 20
DB 645 GTAGAAAACCTTAAAGCAGCAGCCCTTGTCTCTGAGCTGCAAGAAAGATTAACCACTTG 704
QY 21 AsnPheAsnLysAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly 40
DB 705 AACTTCTAATAACATGACATATTAATCTGTGGAGCAGCAAGAAATGGTGGTTGGG 764
QY 41 GluValGlnGlyGlnGlyTrpPheProLysSerTyValLysLeuIleSerGly 59
DB 765 GAGGTGATGAGGAGAGAGATGGTTCCCAATCTTATGTCAAGATCATTCCTGGG 821

RESULT 2
US-09-833-381-1294/c
Sequence 1294, Application US/0983381
Patent No. 6672186
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
CURRENT FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1294
LENGTH: 970
TYPE: DNA
ORGANISM: Homo sapiens
US-09-833-381-1294

Alignment Scores:
Pred. No.: 2,22e-10
Score: 129.50
Percent Similarity: 74.00%
Best Local Similarity: 42.00%
Query Match: 39.48%
DB: 4
Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-833-381-1294 (1-970)

QY 7 GlnAlaLeuTyProTPrpArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsp 26
DB 376 AAGGCTTGATTTCTTTCAAGCCAGCAAGATGATGATTGGAAATTTGAAAAGGGTGAC 317
QY 27 ValIleThrValLeuGlnGlnGlnAspMetTrpTrpPheGlyGlnValGlnGlyGln 45
DB 316 ATTGTGATTATACACGAGAAAAGAAAGAGAGATGCTGTTGATCTTTCAATGGGAAA 257
QY 46 LysGlyTrpPheProLysSerTyValLys 55
```

```
DB 256 AAGGCCATTCTCGCCGCTTATGTGAG 227

RESULT 3
US-09-620-312D-676
Sequence 676, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyen
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Weinman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunging
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: 784C1P2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc fl_genes Version 1.0
SEQ ID NO 676
LENGTH: 2955
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (169)..(2418)
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(2955)
OTHER INFORMATION: n = a, t, c or g
US-09-620-312D-676

Alignment Scores:
Pred. No.: 4,24e-09
Score: 125.50
Percent Similarity: 66.67%
Best Local Similarity: 40.74%
Query Match: 38.26%
DB: 4
Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-620-312D-676 (1-2955)

QY 5 GlnAlaGlnAlaLeuTyProTPrpArgAlaLysLysAspAsnHisLeuAsnPheAsnLys 24
DB 1618 CGAGCCAAAGGCCCTTGAGCTTGAACGGCAGCAGCAGCAGCTGCTGCGCAAG 1677
QY 25 AsnAspValIleThrValLeuGlnGlnGlnAspMetTrpTrpPheGlyGlnValGln 43
DB 1678 AACGACATCATCAACAATCGTGTCTCAGAGAGACGAGCTGCTGGGGGAGACTCAAC 1737
QY 44 GluGlnLysGlyTrpPheProLysSerTyValLysLeuIle 57
DB 1738 GCGCTGCGAGCGCTGCTTCCAGCCCAAGTCTGTGAAATCTCCTG 1779

RESULT 4
US-09-620-312D-675
Sequence 675, Application US/09620312D
```

```
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Keundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aildong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yundong
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pc_fl_genes Version 1.0
/ SEQ ID NO 675
/ LENGTH: 3039
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (169)..(2502)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(3039)
/ OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-675

Alignment Scores:
Pred. No.: 4,41e-09 Length: 3039
Score: 125.50 Matches: 22
Percent Similarity: 66.67% Conservative: 14
Best Local Similarity: 40.74% Mismatches: 17
Query Match: 38.26% Indels: 1
DB: 4 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-620-312D-675 (1-3039)
QY 5 GlnAlaGlnAlaLeuTyrProTTPArgAlaLysLysAspAsnHisLeuAsnPhenAlaLys 24
Db 1702 CGAGCCAAAGCCCTGCTGAGCTTTGAGCGGACGACGACGACGAGCTTGCGCTTCCGCAAG 1761
QY 25 AsnAspValIleThrValLeuGlnGlnAsp--MetTrpTrpPheGlyGluValGln 43
Db 1762 AACGACATCATCATCATCTGCTCTCAGAGGACGACGACCTGCTGCGGAGGCTCTCAAC 1821
QY 44 GlyAlaLysGlyTyrPheProLysSerTyrValLysLeuIle 57
Db 1822 GGCCTCGAGGCTGCTTCCAGCGCAAGTTCGTGAGAGTCTCTG 1863

RESULT 5
US-10-164-595-17
/ Sequence 21, Application US/10164595
/ Patent No. 6657054
/ GENERAL INFORMATION:
/ APPLICANT: Origene Technologies, Inc
/ TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
/ FILE REFERENCE: IU 103 R1
```

```
/ CURRENT APPLICATION NUMBER: US/10/164,595
/ CURRENT FILING DATE: 2002-06-10
/ NUMBER OF SEQ ID NOS: 80
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 17
/ LENGTH: 5325
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (264)..(3473)
/ OTHER INFORMATION:
US-10-164-595-17

Alignment Scores:
Pred. No.: 1.36e-08 Length: 5325
Score: 124.50 Matches: 23
Percent Similarity: 66.67% Conservative: 15
Best Local Similarity: 40.35% Mismatches: 18
Query Match: 37.96% Indels: 1
DB: 4 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-10-164-595-17 (1-5325)
QY 2 GluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaLysLysAspAsnHisLeuAsn 21
Db 3294 AAGGGAGAGAGCCCAAGCCCTTATATGATTTCCGAGGAGATCAAGATGAATTCCTCC 3353
QY 22 PheAsnLysAsnAspValIleThrValLeuGlnGln--GlnAspMetTrpTrpPheGly 40
Db 3354 TTCAGGCTCGAGATATATATACAGAGCTGGAATCTGTAGATCATACATCGAGATGGA 3413
QY 41 GluValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuIle 57
Db 3414 GAACCTATGCGAAATCTGGAATATTTCCCAAAAACATACATACAGTTTCTA 3464

RESULT 6
US-10-164-595-21
/ Sequence 21, Application US/10164595
/ Patent No. 6657054
/ GENERAL INFORMATION:
/ APPLICANT: Origene Technologies, Inc
/ TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
/ FILE REFERENCE: IU 103 R1
/ CURRENT APPLICATION NUMBER: US/10/164,595
/ CURRENT FILING DATE: 2002-06-10
/ NUMBER OF SEQ ID NOS: 80
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 21
/ LENGTH: 5334
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (264)..(3482)
/ OTHER INFORMATION:
US-10-164-595-21

Alignment Scores:
Pred. No.: 1.36e-08 Length: 5334
Score: 124.50 Matches: 23
Percent Similarity: 66.67% Conservative: 15
Best Local Similarity: 40.35% Mismatches: 18
Query Match: 37.96% Indels: 1
DB: 4 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-10-164-595-21 (1-5334)
QY 2 GluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaLysLysAspAsnHisLeuAsn 21
Db 3303 AAGGGAGAGAGCCCAAGCCCTTATATGATTTCCGAGGAGATCAAGATGAATTCCTCC 3362
QY 22 PheAsnLysAsnAspValIleThrValLeuGlnGln--GlnAspMetTrpTrpPheGly 40
```



```

REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1296:
SEQUENCE CHARACTERISTICS:
LENGTH: 1968 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g32054
US-09-023-655-1296

Alignment Scores:
Pred. NO.: 2e-08 Length: 1968
Score: 119.50 Matches: 24
Percent Similarity: 62.50% Conservative: 11
Best Local Similarity: 42.86% Mismatches: 20
Query Match: 36.43% Indels: 1
Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) * US-09-023-655-1296 (1-1968)
QY 3 G1yLeuGlnAlaGlnAlaLeuTyPProTyrArg1AlaYsIySaSPaEnHIsLeuAsnPhe 22
Db 1330 GGGATCTGAGCTGTGGCTCTATATGATTAACCAAGAGAGGAGAGTGAACCTTCCCTTT 1389
QY 23 AsnIySaSPaEnPValIleThValLeuGlnIuGlnAsnMet---TyrTrpPheGlyGlu 41
Db 1390 GATCCGAGCGACGATPACTGACATTAAGATGGTGGACGAGGGGCTGTGGCGGGAGCT 1449
QY 42 ValGlnGlyGlnIySgIyTrpPheProIySseTyrValIySleuIle 57
Db 1450 TGGCATGGCCACTTTGGACTCTTCCCTGCAATTATATGTCAGCTTCGT 1497

RESULT 12
US-08-630-915A-13
: Sequence 13, Application US/08630915A
: Patent No. 6309820
: GENERAL INFORMATION:
: APPLICANT: SPARKS, Andrew B.
: APPLICANT: HOFFMAN, No. 6309820h
: APPLICANT: KAY, Brian K.
: APPLICANT: FOWLES, Dana M.
: APPLICANT: MCCONNELL, Stephen D.
: TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
: TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
: NUMBER OF INVENTION: USING SAME
: NUMBER OF SEQUENCES: 227
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/630,915A
: FILING DATE: 03-APR-1996
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Mistrock, S. Leslie
: REGISTRATION NUMBER: 18,872
: REFERENCE/DOCKET NUMBER: 1101-174
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090

```

```

TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1867 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-915A-13

Alignment Scores:
Pred. No.: 3.76e-08 Length: 1867
Score: 117.50 Matches: 23
Percent Similarity: 65.45% Conservative: 13
Best Local Similarity: 41.82% Mismatches: 18
Query Match: 35.82% Indels: 1
Gaps: 1

US-09-720-934-2_COPY 908_966 (1-59) x US-08-630-915A-13 (1-1867)

Qy 3 GllleuglnalaglnalaLeuTyPrOTPrargalalulslvsaspasnhileuandsbhe 22
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Dh 1598 GGATCAACAGCCATCGCCCTGTATGACTACACAGGCTGCTGGCGATGATGATCCTCTTT 1657
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 23 AsnlvsasnaapvalllleThValleuGluGln---GlnaapmetTyTrPrPhelgylu 41
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Dh 1658 GACCTTAGTACATCATCATCCACACATGAGATTGATTCAGATGGCTGGTGGCGTGGGCTG 1717
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 42 ValGnlglyGlnslgLYTrPrPhaProLysseerTYValylslu 56
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Dh 1718 TGCAAGGCGACATACGGGCTCTTCCACACCAAGATATGAGAGCTG 1762
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

RESULT 13
US-08-630-915A-39
; Sequence 39, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mastrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 181.007 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_908_966
Perfect score: 328
Sequence: 1 VEGIQAQALYPMRAKKDNHL.....GEVGGQGMFKSYVKLISG 59

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Delop 6.0, Delext 7.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+g2n.model -DEV=xlp
-Q/cg2_1/USFTO.spool_p/US09720994/runat_30062004_064541_13481/app_query.fasta_1.1386
-DB=Published Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MARRIX-biosum62
-TRANS-human40.ccd -LIST=45 -DOCCALIGN=200 -THR SCORE=Pct -THR MAX=100
-THR MN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09720994_@CGN_1.1.1500_@runat_30062004_064541_13481
-NPU=6 -ICPU=3 -NO MAP -LARGEDUERY -NEG SCORES=0 -WAIT -DSBHOOK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database : Published Applications_NA:*

1:	/cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cg2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3:	/cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cg2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cg2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10:	/cg2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11:	/cg2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12:	/cg2_6/ptodata/1/pubpna/US09D_NEW_PUB.seq:*
13:	/cg2_6/ptodata/1/pubpna/US09E_NEW_PUB.seq:*
14:	/cg2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15:	/cg2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16:	/cg2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17:	/cg2_6/ptodata/1/pubpna/US10D_NEW_PUB.seq:*
18:	/cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19:	/cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	251	76.5	206	9	US-09-864-761-30400	Sequence 30400, A
2	251	76.5	458	9	US-09-864-761-13834	Sequence 13834, A
3	251	76.5	2873	9	US-09-879-957-193	Sequence 193, App
4	251	76.5	3746	11	US-09-764-875-176	Sequence 176, App
5	251	76.5	4210	9	US-09-764-868-125	Sequence 125, App
6	251	76.5	5828	13	US-10-398-885A-15	Sequence 15, App1
7	134.5	41.0	2067	16	US-10-264-049-887	Sequence 887, App
8	129.5	39.5	2539	9	US-09-833-381-1294	Sequence 1294, App
9	129.5	39.5	2539	16	US-10-104-047-1655	Sequence 1655, App
C 10	125.5	38.3	2803	11	US-09-764-868-115	Sequence 115, App
C 11	125.5	38.3	2955	15	US-10-037-270-676	Sequence 324, App
C 12	125.5	38.3	2955	16	US-10-117-722-675	Sequence 676, App
13	125.5	38.3	2955	15	US-10-117-722-675	Sequence 676, App
14	125.5	38.3	3039	16	US-10-037-270-676	Sequence 675, App
15	125.5	38.3	3039	16	US-10-117-722-675	Sequence 675, App
16	124.5	38.0	389	10	US-09-918-995-23920	Sequence 23920, A
17	124.5	38.0	2688	11	US-09-764-875-198	Sequence 198, App
18	123.5	37.7	581	10	US-09-918-995-14807	Sequence 14907, A
19	121.5	37.0	4668	13	US-10-087-192-1205	Sequence 1205, App
20	121.5	37.0	4760	15	US-10-240-965-1105	Sequence 105, App
21	120.5	36.7	2003	9	US-09-879-957-125	Sequence 25, App1
22	119.5	36.4	1416	16	US-10-388-934-99	Sequence 99, App1
23	119.5	36.4	1968	9	US-09-880-107-3711	Sequence 3711, App
24	119.5	36.4	1968	13	US-10-342-887-1141	Sequence 1141, App
25	119.5	36.4	1968	13	US-10-172-118-1141	Sequence 1141, App
26	119.5	36.4	1968	16	US-10-133-937-91	Sequence 91, App1
27	119.5	36.4	1968	16	US-10-159-563-91	Sequence 91, App1
28	119.5	36.4	1968	17	US-10-641-643-1296	Sequence 1296, App
29	119.5	36.4	4666	16	US-10-159-563-299	Sequence 299, App
30	117.5	35.8	1867	9	US-09-879-957-13	Sequence 13, App1
31	116.5	35.5	747	9	US-09-879-957-13	Sequence 13, App1
32	116.5	35.5	3248	15	US-10-171-581-185	Sequence 185, App
33	116.5	35.5	4053	13	US-10-342-887-1882	Sequence 1882, App
34	116.5	35.5	4053	13	US-10-172-118-1882	Sequence 1882, App
35	115.5	35.2	2278	9	US-09-764-868-131	Sequence 131, App
36	115.5	35.2	2278	11	US-09-764-875-494	Sequence 494, App
37	114.5	34.9	2967	15	US-10-207-655-197	Sequence 197, App
38	113.5	34.6	386	13	US-10-085-783A-37196	Sequence 37196, A
39	113.5	34.6	386	16	US-10-242-535A-37196	Sequence 37196, A
40	113.5	34.6	400	13	US-10-085-783A-22639	Sequence 22639, A
41	113.5	34.6	400	16	US-10-242-535A-22639	Sequence 22639, A
42	113.5	34.6	1215	17	US-10-648-593-61	Sequence 61, App1
43	113.5	34.6	3143	14	US-10-144-621-1	Sequence 1, App1
44	113.5	34.6	3348	13	US-10-342-887-187	Sequence 187, App
45	113.5	34.6	3348	13	US-10-172-118-187	Sequence 187, App

ALIGNMENTS

RESULT 1
US-09-864-761-30400
Sequence 30400, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Weisheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359

```
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 30400
;; LENGTH: 206
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC008073.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
;; OTHER INFORMATION: NT HIT: AF182198.1, EVALUE 1.00e-112
;; OTHER INFORMATION: SWISSPROT HIT: Q15811, EVALUE 3.00e-23
;; OTHER INFORMATION: EST_HUMAN HIT: AW505025.1, EVALUE 1.00e-112
US-09-864-761-30400

Alignment Scores:
Pred. No.: 2,46e-28 Length: 206
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x US-09-864-761-30400 (1-206)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTPaGAlaLysAspAsnHisLeu 20
DB 10 GTAGAAAACCTTAAGACACAGGCCCTTGTCTGCTGACTGCAAAAGATACCACTTG 69
QY 21 AsnPhaenLysAsnAspValIleThrValLeuGlnGlnGlnAspMetTrpTirPheGly 40
DB 70 AACCTTCGAAACATGACATTTACTGTCTTGAGCAGCAAGAAATTGGTGTGTTGGG 129
QY 41 GluValGlnGlyGlnLysGlyTirPheProLysSerTyTValLysLeuLleSerGly 59
DB 130 GAGGTGATGAGGAGAGAGAGATGGTTCCCAATCTTAATGTCAGATCAATTCCTGG 186

RESULT 2
US-09-864-761-13834
; Sequence 13834, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
```

```
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 13834
;; LENGTH: 458
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC008073.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
US-09-864-761-13834

Alignment Scores:
Pred. No.: 6,86e-28 Length: 458
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x US-09-864-761-13834 (1-458)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTPaGAlaLysAspAsnHisLeu 20
DB 247 GTAGAAAACCTTAAGACACAGGCCCTTGTCTGCTGACTGCAAAAGATACCACTTG 306
QY 21 AsnPhaenLysAsnAspValIleThrValLeuGlnGlnGlnAspMetTrpTirPheGly 40
```

Db 307 AACTTCAAAACATGACATTATTACTGTCTTGAGCAGCAAGAAATGGTGGTGGG 366

Qy 41 GluValGlnGlyGlnLysGlyTTPheProLysSerTyrrValLysLeuIleSerGly 59

Db 367 GAGGTGCATGAGAGAGAGATGGTTCCTCCAAATCTTATGTCAAGATCATTCTCTGGG 423

RESULT 3

US-09-879-957-193
; Sequence 193, Application US/09879957
; Patent No. US20020034755A1
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; HOFFMAN, No. US20020034755A1h
; KAY, Brian K.
; FOWLES, Dana M.
; MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/879,957
; FILING DATE: 13-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,915
; FILING DATE: 03-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2873 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 193:
US-09-879-957-193

Alignment Scores:

Pred. No.: 7,21e-27 Length: 2873
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x US-09-879-957-193 (1-2873)

Qy 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrrProTTPArgAlaLysLysAspAsnHisLeu 20

Db 645 GTAGAAACTTAATAAACACAGGCCCTTTGTTCTGACACTGCAGAAAGAAATTAACCACTTG 704

Qy 21 AasnPheAsnLysAsnAspValIleThrValLeuGlnGlnAspMetTTPTPheGly 40

Db 705 AACTTCAAAACATGACATTATTACTGTCTTGAGCAGCAAGAAATGGTGGTGGG 764

Qy 41 GluValGlnGlyGlnLysGlyTTPheProLysSerTyrrValLysLeuIleSerGly 59

Db 765 GAGGTGCATGAGAGAGAGATGGTTCCTCCAAATCTTATGTCAAGATCATTCTCTGGG 821

RESULT 4

US-09-764-875-176
; Sequence 176, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 176
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-875-176

Alignment Scores:

Pred. No.: 1.01e-26 Length: 3746
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
DB: 11 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x US-09-764-875-176 (1-3746)

Qy 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrrProTTPArgAlaLysLysAspAsnHisLeu 20

Db 571 GTAGAAACTTAATAAACACAGGCCCTTTGTTCTGACACTGCAGAAAGAAATTAACCACTTG 630

Qy 21 AasnPheAsnLysAsnAspValIleThrValLeuGlnGlnAspMetTTPTPheGly 40

Db 631 AACTTCAAAACATGACATTATTACTGTCTTGAGCAGCAAGAAATTAACCACTTG 690

Qy 41 GluValGlnGlyGlnLysGlyTTPheProLysSerTyrrValLysLeuIleSerGly 59

Db 691 GAGGTGCATGAGAGAGAGATGGTTCCTCCAAATCTTATGTCAAGATCATTCTCTGGG 747

RESULT 5

US-09-764-868-125
; Sequence 125, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 125
; LENGTH: 4210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-125

Alignment Scores:

Pred. No.: 1.18e-26 Length: 4210
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
DB: 9 Gaps: 0

```
US-09-720-934-2_COPY_908_966 (1-59) x US-09-764-868-125 (1-4210)
Qy 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTPrArgAlaLysAspAsnHisLeu 20
Db 1022 GTAGAAACTTAAGACAGGCGCTTGTCTCGACTGCAAGAAAGATTAACCACTTG 1081
Qy 21 AsnPhaenLysAsnAspValIleThyValLeuGluGlnGlnAspMetTTrpPhgGly 40
Db 1082 AACTTCTCAAAACATGACATTTACTGTCTTGAGACAGCAAGAAATTGGTGGTTGGG 1141
Qy 41 GluValGlnGlyGlnLysGlyTTrpPhProLysSerTyValLysLeuIleSerGly 59
Db 1142 GAGGTGATGAGAGAGAGATGGTTTCCCAATCTTATGTCAGATCATTCCTGGG 1198

RESULT 6
US-10-398-885A-15
; Sequence 15, Application US/10398885A
; Publication No. US20040053282A1
; GENERAL INFORMATION:
; APPLICANT: Sugita, Yuji
; APPLICANT: Hashida, Ryoichi
; APPLICANT: Ogawa, Kaoru
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Obayashi, Masaya
; APPLICANT: Saito, Hirohisa
; APPLICANT: Takahashi, Eiki
; TITLE OF INVENTION: Method of Testing For Allergic Diseases
; FILE REFERENCE: SHIMIZU-07907
; CURRENT APPLICATION NUMBER: US/10/398, 885A
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: PCT/JP01/08937
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: JP 2000-314093
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 5828
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: CDS
; LOCATION: (7)..(5052)
; OTHER INFORMATION:
US-10-398-885A-15

Alignment Scores:
Pred. No.: 1.79e-26 Length: 5828
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
Gaps: 0
DB: 13

US-09-720-934-2_COPY_908_966 (1-59) x US-10-398-885A-15 (1-5828)
Qy 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTPrArgAlaLysAspAsnHisLeu 20
Db 2650 GTAGAAACTTAAGACAGGCGCTTGTCTCGACTGCAAGAAAGATTAACCACTTG 2709
Qy 21 AsnPhaenLysAsnAspValIleThyValLeuGluGlnGlnAspMetTTrpPhgGly 40
Db 2710 AACTTCTCAAAACATGACATTTACTGTCTTGAGACAGCAAGAAATTGGTGGTTGGG 2769
Qy 41 GluValGlnGlyGlnLysGlyTTrpPhProLysSerTyValLysLeuIleSerGly 59
Db 2770 GAGGTGATGAGAGAGAGATGGTTTCCCAATCTTATGTCAGATCATTCCTGGG 2826

RESULT 7
US-10-264-049-887
; Sequence 887, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Blise et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1133P1
; CURRENT APPLICATION NUMBER: US/10/264, 049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 887
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2058)..(2058)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2063)..(2063)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-887

Alignment Scores:
Pred. No.: 1.06e-09 Length: 2067
Score: 134.50 Matches: 26
Percent Similarity: 64.15% Conservative: 8
Best Local Similarity: 49.06% Mismatches: 18
Query Match: 41.01% Indels: 1
Gaps: 1
DB: 16

US-09-720-934-2_COPY_908_966 (1-59) x US-10-264-049-887 (1-2067)
Qy 5 GlnAlaGlnAlaLeuTyProTPrArgAlaLysAspAsnHisLeuAsnPhaenLys 24
Db 308 CAGGTGATGAGATGATGACATACACCGCGCATGATGACCTGCTCAACAAG 367
Qy 25 AsnAspValIleThyValLeuGluGlnGlnAspMet---TTrpPhgGlyGluValGln 43
Db 368 GAGCCAGATCATCAACGTCTCTCAACAAGAGAGACCTGACTGTGGAAGAGATCAAT 427
Qy 44 GlyGlnLysGlyTTrpPhProLysSerTyValLysLeu 56
Db 428 GACCAAGTGGGAGCTTCTCCATCCATTAATGTGAAGCTG 466

RESULT 8
US-09-833-381-1294/c
; Sequence 1294, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1294
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1294
```

Alignment Scores:

Pred. No.: 2.22e-09 Length: 970
 Score: 129.50 Matches: 21
 Percent Similarity: 74.00% Conservative: 16
 Best Local Similarity: 42.00% Mismatches: 12
 Query Match: 39.48% Indels: 1
 DB: 9 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-833-381-1294 (1-970)

QY 7 GlnAlaLeuTyRProTTPaRgAlaLysLysAspAsnHisLeuAsnPhenAsnLysAsnAsp 26

Db 376 AAGGCGCTTGTATCTCTTCAAGCCAGCAAGATGATGATGATTTGGAAAAGGTCAC 317

QY 27 ValIleThrValIleGlu--GlnGlnAspMetTTPheGlyGluValGlnGlyGln 45

Db 316 ATTGTGATTATACACGAGAAAAAGAAAGAGATGGTGGTTGGATCTTTGAATGGGAAA 257

QY 46 LysGlyTTPheProLysSerTyRValLys 55

Db 256 AAGGCCATTCTTCTCGCGCTTATGTGAG 227

RESULT 9

US-10-104-047-1655
 ; Sequence 1655, Application US/10104047
 ; Publication No. US20030236392A1
 ; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: NO. US20030236392A1e1 full length cDNA
 ; FILE REFERENCE: H1-A0105
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; PRIOR FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:

; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1655
 ; LENGTH: 2539
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-104-047-1655

Alignment Scores:

Pred. No.: 7.63e-09 Length: 2539
 Score: 129.50 Matches: 21
 Percent Similarity: 74.00% Conservative: 16
 Best Local Similarity: 42.00% Mismatches: 12
 Query Match: 39.48% Indels: 1
 DB: 16 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-10-104-047-1655 (1-2539)

QY 7 GlnAlaLeuTyRProTTPaRgAlaLysLysAspAsnHisLeuAsnPhenAsnLysAsnAsp 26

Db 2257 AAGGCGCTTGTATCTTCAAGCCAGCAAGATGATGATGATTTGGAAAAGGTCAC 2216

QY 27 ValIleThrValIleGlu--GlnGlnAspMetTTPheGlyGluValGlnGlyGln 45

Db 2317 ATTGTGATTATACACGAGAAAAAGAAAGAGATGGTGGTTGGATCTTTGAATGGGAAA 2376

QY 46 LysGlyTTPheProLysSerTyRValLys 55

Db 2377 AAGGCCATTCTTCTCGCGCTTATGTGAG 2406

RESULT 10

US-09-764-868-115/c
 ; Sequence 115, Application US/09764868
 ; Patent No. US2002016871A1
 ; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232

CURRENT APPLICATION NUMBER: US/09/764,868

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1510

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 115

; LENGTH: 2803

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-868-115

Alignment Scores:

Pred. No.: 3.42e-08 Length: 2803
 Score: 125.50 Matches: 22
 Percent Similarity: 66.67% Conservative: 14
 Best Local Similarity: 40.74% Mismatches: 17
 Query Match: 38.26% Indels: 1
 DB: 9 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-764-868-115 (1-2803)

QY 5 GlnAlaLeuTyRProTTPaRgAlaLysLysAspAsnHisLeuAsnPhenAsnLys 24

Db 1186 CGAGCCAGGCGCTGCTGAGCTTTGAGCGCAGCAGCAGAGCTGGCTTCGCAAG 1127

QY 25 AsnAspValIleThrValIleGluGlnGlnAsp--MetTTPheGlyGluValGln 43

Db 1126 AAGCAGATCATCAATCATGTCTCAGAGAGCAGCAGCTGCTGGGGGAGACTCAAC 1067

QY 44 GlyGlnLysGlyTTPheProLysSerTyRValLysLeu 57

Db 1066 GGCTGCGAGGCTGGTTTCCAGCCAACTGTGGAAGTCTCTG 1025

RESULT 11

US-09-764-875-324/c
 ; Sequence 324, Application US/09764875
 ; Publication No. US20040018969A1
 ; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: P202
 ; CURRENT APPLICATION NUMBER: US/09/764,875
 ; PRIOR FILING DATE: 2001-01-17
 ; PRIOR APPLICATION data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1249

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 324

; LENGTH: 2803

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-875-324

Alignment Scores:

Pred. No.: 3.42e-08 Length: 2803
 Score: 125.50 Matches: 22
 Percent Similarity: 66.67% Conservative: 14
 Best Local Similarity: 40.74% Mismatches: 17
 Query Match: 38.26% Indels: 1
 DB: 11 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-764-875-324 (1-2803)

QY 5 GlnAlaLeuTyRProTTPaRgAlaLysLysAspAsnHisLeuAsnPhenAsnLys 24

Db 1186 CGAGCCAGGCGCTGCTGAGCTTTGAGCGCAGCAGCAGAGCTGGCTTCGCAAG 1127

QY 25 AsnAspValIleThrValIleGluGlnGlnAsp--MetTTPheGlyGluValGln 43

Db 1126 AAGCAGATCATCAATCATGTCTCAGAGAGCAGCAGCTGCTGGGGGAGACTCAAC 1067

QY 44 GlyGlnLysGlyTTPheProLysSerTyRValLysLeu 57

Db 1066 GGCTGCGAGGCTGGTTTCCAGCCAACTGTGGAAGTCTCTG 1025


```
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1e1 Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 675
; LENGTH: 3039
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (169)..(2502)
; NAME/KEY: misc_feature
; LOCATION: (1)..(3039)
; OTHER INFORMATION: n = a,t,c or g
US-10-037-270-675
```

```
Alignment Scores:
Pred. No.: 3,79e-08 Length: 3039
Score: 125.50 Matches: 22
Percent Similarity: 66.67% Conservative: 14
Best Local Similarity: 40.74% Mismatches: 17
Query Match: 38.26% Indels: 1
DB: 15 Gaps: 1
```

US-09-720-934-2_COPY_908_966 (1-59) x US-10-037-270-675 (1-3039)

```
QY 5 GlnAlGlnAlaLeuTyRProTPArgAlaLysAspAsnHisLeuAsnPhaAnlys 24
   ::::::::::::::::::::
Db 1702 CGAGCCAGAGCCCTGCTGAGCTTTGAGCGGACAGACGACGAGCTGGGCTTCCGCAAG 1761
   |||||
QY 25 AsnAspValIleThrValLeuGlnGlnAsp---MetTrpTrpPheGlyGluValGln 43
   |||||
Db 1762 AACGACATCATCAATCGTCTCAGAAAGACGACGACTGCTGGGTGGGGAGCTCAAC 1821
   |||||
QY 44 GlyGlnLysGlyTrpPheProlysSerTyRValLysLeuIle 57
   |||||
Db 1822 GGCTGTGAGGCTGTGTTTCCAGCCAGACTTGTGGAAGTCTCTG 1863
   |||||
```

RESULT 15

```
US-10-117-722-675
; Sequence 675, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1e1 Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 675
; LENGTH: 3039
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (169)..(2502)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(3039)
; OTHER INFORMATION: n = a,t,c or g
US-10-117-722-675
```

```
Alignment Scores:
Pred. No.: 3,79e-08 Length: 3039
Score: 125.50 Matches: 22
Percent Similarity: 66.67% Conservative: 14
Best Local Similarity: 40.74% Mismatches: 17
Query Match: 38.26% Indels: 1
DB: 16 Gaps: 1
```

US-09-720-934-2_COPY_908_966 (1-59) x US-10-117-722-675 (1-3039)

```
QY 5 GlnAlGlnAlaLeuTyRProTPArgAlaLysAspAsnHisLeuAsnPhaAnlys 24
   ::::::::::::::::::::
Db 1702 CGAGCCAGAGCCCTGCTGAGCTTTGAGCGGACAGACGACGAGCTGGGCTTCCGCAAG 1761
   |||||
QY 25 AsnAspValIleThrValLeuGlnGlnAsp---MetTrpTrpPheGlyGluValGln 43
   |||||
Db 1762 AACGACATCATCAATCGTCTCAGAAAGACGACGACTGCTGGGTGGGGAGCTCAAC 1821
   |||||
QY 44 GlyGlnLysGlyTrpPheProlysSerTyRValLysLeuIle 57
   |||||
Db 1822 GGCTGTGAGGCTGTGTTTCCAGCCAGACTTGTGGAAGTCTCTG 1863
   |||||
```

Search completed: July 1, 2004, 20:16:02
Job time : 188.007 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 1774.84 Seconds
(without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_908_966
Perfect score: 328
Sequence: 1 VEGIQAOALYPMWRAKKDNHL.....GEVQGXKWPXKSTVKLISG 59

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODER=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.P/US09720934/rumat.30062004.064540.13442/app_query.fasta_1.1386
-DB=BST -QFMT=fastp -SUFFIX=rest -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -LIST=45
-UNITS=bites -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45
-DOCAI=GN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -AUGN=15 -MODE=LOCAL
-OUTFMT=ext -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934@cgn 1.1 12421 @rumat.30062004.064540.13442 -NCPU=6 -ICPU=3
-NO WMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em.esbba:*
2: em.esthm:*
3: em.estin:*
4: em.estmu:*
5: em.estov:*
6: em.estpl:*
7: em.estro:*
8: em.htc:*
9: gb.est1:*
10: gb.est2:*
11: gb.htc:*
12: gb.est3:*
13: gb.est4:*
14: gb.est5:*
15: em.estfun:*
16: em.estcom:*
17: em.gss.hum:*
18: em.gss.inu:*
19: em.gss.pln:*
20: em.gss.vrt:*
21: em.gss.fun:*
22: em.gss.mus:*
23: em.gss.man:*
24: em.gss.pio:*
25: em.gss.rod:*
26: em.gss.pdg:*
27: em.gss.vbl:*
28: gb.gset1:*

29: gb.gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	328	100.0	527	9	AL120063	AL120063 DKFZp761L192
2	328	100.0	536	10	B1340403	B1340403 365659 MA
3	328	100.0	553	12	AM643494	AM643494 cm30d04.w
4	328	100.0	556	9	AV590689	AV590689 AV590689
5	328	100.0	582	10	BP911573	BP911573 MR4-UTG09
6	328	100.0	651	13	BU272390	BU272390 603509535
7	324	98.8	800	14	CB756509	CB756509 AGENCOURT
8	324	98.8	847	14	CB756047	CB756047 AGENCOURT
9	323	98.5	597	14	CF172865	CF172865 B0915A02-
10	323	98.5	606	14	CF174208	CF174208 B0935F02-
11	323	98.5	738	12	BM944544	BM944544 UI-M-EHDP
12	323	98.5	5385	11	BC062938	BC062938 Mus muscu
13	320	97.6	434	14	CA360651	CA360651 634174 NC
14	310	94.5	690	13	BQ179493	BQ179493 UI-M-EWO-
15	310	94.5	694	14	CB435046	CB435046 611770 MA
16	280	85.4	341	10	AM889319	AM889319 QV4-NT002
17	276	84.1	330	10	BE763158	BE763158 QV4-NT002
18	253	77.1	338	14	R93190	R93190 YC69901.81
19	252	76.8	845	13	BU455276	BU455276 603771924
20	251	76.5	500	10	AM505025	AM505025 UI-HF-BNO
21	251	76.5	961	12	BM466259	BM466259 AGENCOURT
22	250	76.2	481	14	CA886530	CA886530 B0130E05-
23	250	76.2	482	28	CC178744	CC178744 XBA4.8 Bay
24	250	76.2	499	14	CA889924	CA889924 B0156A01-
25	250	76.2	522	12	BU095151	BU095151 B0705151
26	250	76.2	708	12	B1155542	B1155542 602903945
27	249	75.9	465	14	CB715459	CB715459 AMGNNUC:M
28	245	74.7	635	14	CA327260	CA327260 UI-M-FY0-
29	244	74.4	476	10	AM761705	AM761705 UG22d09.Y
30	242	73.8	406	28	CC200213	CC200213 XG015 Bay
31	235	71.6	343	14	CB693264	CB693264 AMGNNUC:C
32	234	71.3	587	9	AV595297	AV595297 AV595297
33	230	70.1	727	14	CF539609	CF539609 UI-M-GVO-
34	223	68.0	650	10	BB656585	BB656585 BB656585
35	212	64.6	397	10	BE702057	BE702057 QV0-NN102
36	208	63.4	760	13	EX879964	EX879964 EX879964
37	202	61.6	521	9	A1303871	A1303871 u157d09.Y
38	202	61.6	653	14	CF136055	CF136055 UI-HF-BNO
39	198	60.4	555	9	AV960622	AV960622 AV960622
40	197	60.1	527	14	CF729468	CF729468 UI-M-HDO-
41	187	60.1	600	14	CF738758	CF738758 UI-M-HDO-
42	184	56.1	668	14	CF763755	CF763755 CBS005445
43	180	54.9	628	13	BQ393010	BQ393010 NISC_mg28
44	179	54.6	492	10	AM206547	AM206547 UI-H-BI-
45	150	45.7	641	28	BZ399663	BZ399663 EINAK63TF

ALIGNMENTS

RESULT 1
AL120063
LOCUS AL120063 527 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp761L192_r1 761 (synonym: hamy2) Homo sapiens cDNA clone
ACCESSION AL120063
VERSION AL120063.1 GI:59259662
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 527)

AUTHORS	Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE	EST (Ottenwaelder, et al.)
JOURNAL	Unpublished (1999)
COMMENT	Contact: MIPS

This clone (DKFZp761L192) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, Germany; Email: clone@rzpd.de. Location/chr3: 15.5 Mb

FEATURES

SOURCE

```

/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="DKFZP761L192"
/ctssue_type="amygdala"
/dev_stage="adult"
/lab_host="DH10B"
/clone_l1b="761" (synonym: hamy2)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

```

Alignment Scores:	
Pred. No.:	5, 82e-36
Score:	328.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	100.00%
DB:	9
Length:	527
Matches:	59
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-720-934-2_COPY_908_966 (1-59) * ALL20063 (1-527)

1 valaiguglyleuinalaaginalaeturypotrlprgalaalsysaspasnhisleu 20
 Dd 306 GTGAGGGGCTACAAGCTCAAGCCCTATATCTTGAGAGCCAAAAAGACAAACCACTTA 36

Dy 21 AsnPhen^aLeuLysAsnAspValIleThrValleugInginAspMetTrpPheGly 40
Dd 366 AATTTTACAAAATGATGCATCACCGTCCTGGAAcAGAcAGAcACTGTGGTGTGGA 425

41 GINVAAGNGIYGILNLSGIYTHPPhProlysserTYrValylsLeuIleSerGly 59
426 GAAGTCAAAGGTCAGAAAGGTTGGTTCGCCAAGTCTTACGTGAACATCATTTACGGG 482

Result 2	LOCUS	BI340403	536 bp	MRNA	linear	EST 30-JUL-2001
BI340403	DEFINITION	365659	MARC 2P1G	Sus scrofa	CDNA 5', mRNA sequence.	
	ACCESION	B1740403				

SOURCE	Sus scrofa (pig)
ORGANISM	Sus scrofa

REFERENCE	Country	Year
1 (bases 1 to 536)		
Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,		
AUTHORS		

TITLE	JOURNAL	YEAR
Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly	Mamm. Genome 13 (8), 475-478 (2002)	2002

MIDDLE 44213789
PUBMED 12226715
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smit@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mnscore 18
and -mismatch 12 options.
GCR. trimmatch

FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 105 row: A column: 4
 Seq primer: ATTTAGTGACACTATAG

```

/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_1b="MARC 2PR3"
/notes="Vector: pCMV.SPORT6, Site_1: NotI, Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

```

Alignment Scores

Pred. No.:	5	96e-36	Length:	53
Score:	328.00		Matches:	59
Percent Similarity:	100.00%		Conservative:	0
Best Local Similarity:	100.00%		Mismatches:	0
Query Match:	100.00%		Indels:	0
DB:	12		Gaps:	0

US-09-720-934-2_COPY_908_966 (1-59) X BI340403 (1-536)

QY 1 ValGIUGlyLeuGlnAlaGlnAlaLeuTyrProTArgAlaLysLysAspAsnHisLeu 20
Dp 322 GTGGAGGGGCTCCAGGCCCAAGCCCTGTATCTTGGAGAGCCAAAAAGACAACCATTTA 38

QY 21 AaNPheasnlYsaasnaSPValIleThrValleuGluGlnGlnaSPmetTriPTyrPheGly 40

Db 382 AATTTTAAcAAaATGATGTcATcACcGTCTTGAGAcAGAcAGAcATGTGTGTTGGA 441

DY 41 GlnValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuIleSerGly 59
DB 442 GAGGTTCAAGGCCAGAGGGTTGGTTCCTAAGTCTACGTGAACCTATTTCAGGG 498

SOURCE	Xenopus laevis (African clawed frog)
ORGANISM	Xenopus laevis

Jr. Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M., Touchman, J.W., Bonaldo, M.F. and Soares, M.B. The NIHES Xenopus maternal EST project: Interim analysis of the first 13,879 ESTs from unfertilized eggs
Journal Gene 267 (1), 71-87 (2001)

MEDLINE	21211403
PUBMED	11311557
COMMENT	Contact: Perry J. Blackshear Office of Clinical Research and Laboratory of Signal Transduction National Institute of Environmental Health Sciences A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,

USA
Tel: 919 541-4899
Fax: 919 541-4571
Email: black009@niehs.nih.gov
Clone is available through Research Genetics, Inc., 2130 Memorial Parkway, Huntsville, AL 35901
phone 800-533-4363 ext.cdn, fax 256-536-9016 att:cdna, email cdna@resgen.com
DNA Sequencing and analyses performed by National Institutes of Health Intramural Sequencing Center (NISC).
PCR Primers
FORWARD: TGTAAACGACGGCCAGT
BACKWARD: CAGGAACAGCTATGACC
Plate: 0129 row: D column: 04
Seq primer: T7 primer.
Location/Qualifiers
1. 553
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="PBX0129D04"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/clone_11b="Blackshear/Soares normalized Xenopus egg library"
/note="Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; PolyA-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT73-Pac as described in Bonaldo, M.F., Lemmon, G. and Soares, M.B. 'Normalization and subtraction: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-dt18 primer; double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT73-Pac vector. The library contained approximately 7.2 X 10⁵ recombinants, with average insert sizes of 1-1.5 kb."

ORIGIN

Alignment Scores:
Pred. No.: 6,24e-36 Length: 553
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AW643494 (1-553)

QY 1 ValG1G1YleuG1n1aG1n1a1eU1YrPr0T1P1rG1A1a1y1s1yS1p1A1n1s1t1eU 20
DB 174 G1TGAAGGCTTACGACCAAGCCTTGTATCTTGAAGCAAGAAGCAACCATCTT 233
QY 21 AasnPhaSn1yS1aSn1pV1l1e1n1rV1a1l1eU1G1n1G1n1aSp1e1t1P1T1P1P1eG1Y 40
DB 234 AATTTTAAACAAAATGATGATATCAAGTCTTGAACAGAGATATGTGGTGGG 293
QY 41 GluValG1nG1y1n1yS1G1YrTp1P1eP1r01yS1e1T1YrV1a1l1yS1e1u1l1eS1e1rG1Y 59
DB 294 GAAGTTCAGGTCAAGGAGGAGGCTGTTCCCAATCCATGTAAAGCTTATATCCGGT 350

RESULT 4
LOCUS AV590689 556 bp mRNA linear EST 27-NOV-2001
DEFINITION AV590689 Bos taurus brain fetus Bos taurus cDNA clone F1BR014A10
5', mRNA sequence.
ACCESSION AV590689
VERSION AV590689.1 GI:9701682
KEYWORDS EST.
SOURCE Bos taurus (cow)

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 556)
Takesuga, A., Hirotsune, S., Itoh, R., Itohzono, A., Suzuki, H., Aso, H. and Sugimoto, Y.
Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
JOURNAL MEDLINE 21570554
PUBMED 11713328
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: Kazusugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers
1. 556
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="F1BR014A10"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_11b="Bos taurus brain fetus"
/note="Vector: pZ1; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"

ORIGIN

Alignment Scores:
Pred. No.: 6,28e-36 Length: 556
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AV590689 (1-556)

QY 1 ValG1G1YleuG1n1aG1n1a1eU1YrPr0T1P1rG1A1a1y1s1yS1p1A1n1s1t1eU 20
DB 55 G1TGAAGGCTTACGACCAAGCCTTGTATCTTGAAGCAAGAAGCAACCATTTA 114
QY 21 AasnPhaSn1yS1aSn1pV1l1e1n1rV1a1l1eU1G1n1G1n1aSp1e1t1P1T1P1P1eG1Y 40
DB 115 AATTTTAAACAAAACCAAGCTTACCGTATCGTGAAGCCCAAGAAGCAACCATTTA 174
QY 41 GluValG1nG1y1n1yS1G1YrTp1P1eP1r01yS1e1T1YrV1a1l1yS1e1u1l1eS1e1rG1Y 59
DB 175 GAAGTTCAGGCTCAAGGAGGAGGCTTCCCAATCCATGTAAGTCAATCTTATTCAGG 231

RESULT 5
LOCUS BF911573 582 bp mRNA linear EST 18-JAN-2001
DEFINITION BF911573 MR4-UT0090-091100-002-e07 UT0090 Homo sapiens cDNA, mRNA sequence.
VERSION BF911573
KEYWORDS BF911573.1 GI:12303031
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE
20202663
10737800
PUBMED

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR4&ct2=MR4-UT0090-091100-002-e07&ct3=2000-11-09&ct4=1)
Seq primer: puc 18 forward
High quality sequence start: 40
High quality sequence stop: 555.
Location/Qualifiers
1. 582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1lb="UT0090"
/note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Alignment Scores:
Pred. No.: 6,71e-36 Length: 582
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x BP911573 (1-582)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTTPArgAlaIySlySaSPaNHisleu 20
Db 334 GTGAGAGGCGCTACAGCTCAAGCCCTATATCTTGAGAGCCAAAAAGACACCACTTA 275

QY 21 AspPheAsnIySaSaSPaVAlIleThValLeuGluGlnGlnAspMetTTPTPheGly 40
Db 274 AATTTTAAACAAAATGATGTCATCACCGTCTGGAACAGACATGCTGTGTTTGA 215

QY 41 GluValGlnGlyGlnIySgIyTTPheProIySeSerTyValIySleuIleSergly 59
Db 214 GAAGTTCAGAGTCAGAAAGGTTGGTCCCAAGCTTACCTGAAACTCATTTTCAAGG 158

RESULT 6
BU272390 651 bp mRNA linear EST 26-NOV-2002
LOCUS 603509535F1 CSROCHN52 Gallus gallus cDNA clone CHEST439f10 5', mRNA
DEFINITION
Sequence.
ACCESSION BU272390
VERSION BU272390.1 GI:25543327
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.
1 (bases 1 to 651)

REFERENCE
AUTHORS
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

JOURNAL MEDLINE
22335534
PUBMED

COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. 651
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, HiseX"
/db_xref="taxon:9031"
/clone="CHEST439f10"
/dev_stage="22"
/lab_host="DH10B"
/clone_1lb="CSROCHN52"
/note="Organ: limbs; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was bluntended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 7.87e-36 Length: 651
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x BU272390 (1-651)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTTPArgAlaIySlySaSPaNHisleu 20
Db 373 GTGAGAGGACTTCACAGCAGCGCTGTATCTTGAGAGCAAAAAAGACACCACTTA 432

QY 21 AspPheAsnIySaSaSPaVAlIleThValLeuGluGlnGlnAspMetTTPTPheGly 40
Db 433 AATTTTAAACAAAATGATGTCATCACGTTTGAAGCAGCAAGATGCTGTGTTTGA 492

QY 41 GluValGlnGlyGlnIySgIyTTPheProIySeSerTyValIySleuIleSergly 59
Db 493 GAAGTTCAGAGCAAAAAGGCTGTTTCCCAAGTCATATGTGAAGCTTATTTCAGG 549

RESULT 7
CB756509/c 800 bp mRNA linear EST 16-MAY-2003
LOCUS AGENCOURT 12983401 NICHD XGC Tad1 Xenopus laevis cDNA clone
DEFINITION
IMAGE:6877279 3', mRNA sequence.
ACCESSION CB756509
VERSION CB756509.1 GI:29836016
KEYWORDS
SOURCE Xenopus laevis (African clawed frog)

ORGANISM	Xenopus laevis
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae; Xenopodinae; Xenopus.
AUTHORS	1 (bases 1 to 800)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Sequenced with vector primer Tissue Procurement: Drs. Donald Brown and Liqun Cai cDNA Library Preparation: CLONTECH cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: L10CM3115 row: k column: 06 High quality sequence stop: 589. Location/Qualifiers 1..800 /organism="Xenopus laevis" /mol_type="mRNA" /db_xref="taxon:8355" /clone="IMAGE:6877279" /dev_stage="metamorphosis stage 53" /clone_lib="NICHD XGC Tadi" /note="Organ: Developing Tadpole; Vector: pDNR-LIB; Site_1: Sfi; Site_2: Sfi; 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCCGCGCATTTATGGCC-3' and 3' adaptor sequence: 5'-ATCCGCGCGCGCGCGCGCATATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.6 kb (range 0.9-3.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
ORIGIN	
Alignment Scores:	
Pred. No.:	3,896-35
Score:	324.00
Percent Similarity:	100.00%
Best Local Similarity:	96.61%
Query Match:	98.78%
DB:	14
US-09-720-934-2_COPY_908_966 (1-59) x CB756509 (1-800)	
QY	1 ValGluGlyLeuGlnAlaGlnAlaLeuYrProTIPATGAlaAlaYlYsAspAsnHiSleu 20
Db	609 GTAGAAGGCTTCACAGCACAAAGCCTTGTAATCCGTGAGAGCAAGAAAGACACACATCTC 550
QY	21 AsnPheAsnLYsAsnAspValIleThrValLeuGlnGlnGlnAspMetTIPTrPheGly 40
Db	549 AATTTTAACAAAATGATATATATACCGGTTCTTGAACAGCAGCATATGTGGGTTTGGG 490
QY	41 GluValGlnGlnGlyGlnGlyGlyTrpPheProLYsSerTyrValYsLeuIleSerGly 59
Db	489 GAAGTTCAAGTCAAAAGGGCGTGGTTTCCCAATCTTATGTAAAGCTTATATGCGGCT 433
RESULT 8	
CB756047/c	
LOCUS	CB756047 847 bp mRNA linear EST 16-MAY-2003
DEFINITION	AGENCOURT_12983387 NICHD_XGC_Tadi Xenopus laevis cDNA clone
ACCESSION	IMAGE:6877279 3', mRNA sequence.
VERSION	CB756047
KEYWORDS	CB756047.1 GI:29835552
SOURCE	EST
ORGANISM	Xenopus laevis (African clawed frog) Xenopus laevis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia: Batrachia; Anura; Mesobatrachia: Pipodae; Pipidae;
Xenopodinae: Xenopus.
1 (bases 1 to 847)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Sequenced with anchor primer
Tissue Procurement: Drs. Donald Brown and Liqun Cai
cDNA Library Preparation: CLONTECH
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
plate: L1CM315 row: k column: 06
High quality sequence start: 2
High quality sequence stop: 746.
Location/Qualifiers
1. 847
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8335"
/clone="IMAGE:6877279"
/dev_stage="metamorphosis stage 53"
/clone_1ib="NICHD_XCC_Tad1"
/note="Organ: Developing Tadpole; Vector: pDNR-LIB;
Site 1: Sfi; Site 2: Sfi; 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGCCCATTTGACC-3' and 3' adaptor sequence:
5'-ATTTAGAGCCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.6 kb
(range 0.9-3.0 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Alignment Scores:

Pred. No.:	4,22e-35	Length:	847
Score:	324.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	96.61%	Mismatches:	0
Query Match:	98.78%	Indels:	0
DB:	14	Gaps:	0

US-93-720-934-2_COPY_908_966 (1-59) x CBF56047 (1-847)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTPRAGAlaSLyASpASHisLeu 20

Db 538 GTAGAGGCTCTTCGAGGACAAGCCCTTGATCCGAGAGCAAGAAAGAACACCACTTC 479

QY 21 AsnheAsnLysAsnaAspValIleThrValLeuGluGlnGlnAspMetTYrTTPPhcGly 40

Db 478 AATTTTAACAAAAATGATATTATACAGGTTCTTAAACGACGAGATATGTGTGTTGGG 419

QY 41 GluValGlnGlyGlnLysGlyTYrTPheProLysSerTYrValLysLeuIleSerGly 59

Db 418 GAAGTCAAGCTCAAAAGGCTGTGGTTTCCCAATCTCATGTAAAGCTATATAGCCGGT 362

RESULT 9

CF172865 597 bp mRNA linear EST 25-JUL-2003

LOCUS B0915A02-5 NTA Mouse Unfertilized Egg cDNA library (long 1) Mus

DEFINITION musculus cDNA clone NTA:B0915A02 IMAGE:30473665 5', mRNA sequence.

ACCESSION CF172865

VERSION CF172865.1 GI:33282414

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 597)
Piao, Y., Ko, N.T., Lam, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE
PUBMED
21429098
11544199
COMMENT
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: B0935 row: A column: 02
Seq primer: M13 Reverse
High quality sequence stop: 597
POLY-A-No.

FEATURES
Source
1. 597
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="tae:0915A02-5"
/db_xref="taxon:10090"
/clone="NIA:B0915A02 IMAGE:30473665"
/dev_stage="Unfertilized Egg"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (long
1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total
RNAs were extracted from a pool of 1488 unfertilized eggs.
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer (Invitrogen):
5'-pGACTAGTCTAGATCGAGCGCCGCTTTT-3',
treated with T4 DNA polymerase, and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker L1-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.5 kb. The library was
constructed by Yulan Piao."

ORIGIN
Alignment Scores:
Pred. No.: 3.55e-35 Length: 597
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.31% Mismatches: 0
Query Match: 98.48% Indels: 0
DB: 14 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-597) * CFI72865 (1-597)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProGlyArgAlaGlyAspAsnHisLeu 20
Db 352 GTGAGAGGGCTACAGCGCAAGCCCTGTATCCCTCGAGACCCCAAAAAGACAACTT 411
QY 21 AasnPhaAsnlySAsnAepValIleThValLeuGluGlnGlnAspMetTPTTPPhaGly 40
Db 412 AATTTTAAACAAGTAGTCATCACCGTTCTGGAACAGCAAGCATGTGTGTTGGA 471
QY 41 GluValGlnGlyGlnGlyTTPPhaPheProLysSerTyValLysLeuIleSerGly 59

DB 472 GAGCTTCAAGTCCAGAGGGTTGGTTCCCAAGTCTTACGTAAACTATTCACGG 528
|||||
RESULT 10
CFI74208 606 bp mRNA linear EST 25-JUL-2003
LOCUS CFI74208
DEFINITION B0935F02-5 NIA Mouse Unfertilized Egg cDNA Library (long 1) Mus
musculus cDNA clone NIA:B0935F02 IMAGE:30475645 5', mRNA sequence.
ACCESSION CFI74208
VERSION CFI74208.1 GI:33283757
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE
1 (bases 1 to 606)
Piao, Y., Ko, N.T., Lam, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE
PUBMED
21429098
11544199
COMMENT
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: B0935 row: F column: 02
Seq primer: M13 Reverse
High quality sequence stop: 606
POLY-A-No.

FEATURES
Source
1. 606
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="tae:0935F02-5"
/db_xref="taxon:10090"
/clone="NIA:B0935F02 IMAGE:30475645"
/dev_stage="Unfertilized Egg"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (long
1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total
RNAs were extracted from a pool of 1488 unfertilized eggs.
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer (Invitrogen):
5'-pGACTAGTCTAGATCGAGCGCCGCTTTT-3',
treated with T4 DNA polymerase, and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker L1-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.5 kb. The library was
constructed by Yulan Piao."

ORIGIN
Alignment Scores:
Pred. No.: 3.62e-35 Length: 606
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1

Best Local Similarity: 98.31% Mismatches: 0
 Query Match: 98.48% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x CF174208 (1-606)

QY 1 ValGluGlyLeuGlnAlaLeuTyrrProTrrArGAlaLysLysAspAsnHisLeu 20
 |||
 DB 217 GTGGAAAGGGCTACAGAGGCAAGCCCTGTATCCTCGAGAGCCAAAAGCAACCACTTA 276
 |||
 QY 21 AsnPhaAsnLysAsnAspValIleThrValleuGluGlnGlnAspMetTrpTrpPheGly 40
 |||
 DB 277 AATTTTACAAAGTACGATCAGCCGTTCTGGAACAGCAACATGTCGTTTGA 336
 |||
 QY 41 GluValGlnGlyGlnLysGlyTyrrPheProLysSerTyrrValLysLeuIleSerGly 59
 |||
 DB 337 GAAGTTCAAGGTGAGAGGGTTGTTCCCAAGTCTTACGTGAACATTCATTTCAGGG 393
 |||

RESULT 11
 BM944544 738 bp mRNA linear EST 14-MAR-2002
 LOCUS UI-M-EH0P-bvr-j-08-0-UI.r1 NIH_BMAP_EH0P Mus musculus cDNA clone
 DEFINITION IMAGE:5695975 5', mRNA sequence.

ACCESSION BM944544
 VERSION BM944544
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS NIH-MGC
 TITLE 1 (bases 1 to 738)
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.liml.gov

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..738
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5695975"
 /tissue_type="whole brain"
 /dev_stage="embryo 18.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EH0P"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCCAGGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

FEATURES

source

Alignment Scores:

Pred. No.: 4.8e-35 Length: 738
 Score: 323.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.31% Mismatches: 0
 Query Match: 98.48% Indels: 0
 DB: 12 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BM944544 (1-738)

QY 1 ValGluGlyLeuGlnAlaLeuTyrrProTrrArGAlaLysLysAspAsnHisLeu 20
 |||
 DB 407 GTGGAAAGGGCTACAGAGGCAAGCCCTGTATCCTCGAGAGCCAAAAGCAACCACTTA 466
 |||
 QY 21 AsnPhaAsnLysAsnAspValIleThrValleuGluGlnGlnAspMetTrpTrpPheGly 40
 |||
 DB 467 AATTTTACAAAGTACGATCAGCCGTTCTGGAACAGCAACATGTCGTTTGA 526
 |||
 QY 41 GluValGlnGlyGlnLysGlyTyrrPheProLysSerTyrrValLysLeuIleSerGly 59
 |||
 DB 527 GAAGTTCAAGGTGAGAGGGTTGTTCCCAAGTCTTACGTGAACATTCATTTCAGGG 583
 |||

RESULT 12
 BC062938 5385 bp mRNA linear HTC 11-DEC-2003
 LOCUS Mus musculus intersectin (SH3 domain protein 1A), mRNA (cDNA clone
 IMAGE:6839463), containing frame-shift errors.

ACCESSION BC062938
 VERSION BC062938.1 GI:38566052
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 1 (bases 1 to 5385)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Dege,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
 Mckernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huiyk,S.W.,
 Vialalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
 Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,D.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butlerfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalys,D.E.,
 Scherch,A., Schein,D.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL PUBLISHED 12477932
 REFERENCE 2 (bases 1 to 5385)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (26-NOV-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT Contact: MGC help desk
 Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
 Thomas L. Casavant.

ORIGIN

Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu, tom-casaevant@uiowa.edu
 Bontaio,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
 Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
 Schaefer,T., Smith,C., Snider,E., Tack,D., Trout,K., Walters,J.,
 Casaevant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: Place: Row: Column: 0
 This clone has the following problem: frame shifted.

FEATURES

source

Location/Qualifiers
 1..5385
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6839463"
 /issue_type="Brain"
 /clone_lib="NIH BMAP_GHO"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"

ORIGIN

Alignment Scores:
 Pred. No.: 8.17e-34 Length: 5385
 Score: 323.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.31% Mismatches: 0
 Query Match: 98.48% Indels: 0
 DB: 11 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BC062938 (1-5385)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTTPArgAlaLysLysAspAsnHisLeu 20

Db 2976 GTGAGAGGGCTACAGCGCAGCCCTGTATCCTGTGAGAGCCAAAGAACACACACTTA 3035

QY 21 AsnPhaSnlyAsnAspValIleThrValLeuGluGlnGlnAspMetTyTrpPheGly 40

Db 3036 AATTTTAAACAAAGTACGACGTATCACTTCTGAAACAGCAACATGTGTGTGTTGGA 3095

QY 41 GluValGlnGlyGlnLysGlyTyTrpPheProLysSerTyValLysLeuIleSerGly 59

Db 3096 GAATTTCAAGTCAAGAGGTTGTGTTCCCAAGCTTACGGAACATCATTTTCAGGG 3152

RESULT 13

CA360651 434 bp mRNA linear EST 06-NOV-2002

LOCUS 634174 NCCMWA 1RT Oncorhynchus mykiss cDNA clone 1RT108K03_A_F02

DEFINITION 5', mRNA sequence.

ACCESSION CA360651.1 GI:24663989

VERSION CA360651.1 GI:24663989

KEYWORDS EST.

SOURCE Oncorhynchus mykiss (rainbow trout)

ORGANISM Oncorhynchus mykiss

REFERENCE 1 (bases 1 to 434)

AUTHORS Raxroad,C.E. and Keele,J.W.

TITLE Sequence analysis of a rainbow trout normalized cDNA library

JOURNAL Unpublished (2002)

COMMENT Contact: Rexroad CE

USDA, ARS, National Center for Cool and Cold Water Aquaculture

11876 Leetown Road, Kearneysville, WV 25430, USA

FEATURES

Location/Qualifiers

1..434
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="1RT108K03_A_F02"
 /issue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="NCCMWA 1RT"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 library made from pooled tissue from brain, gill, liver,
 spleen, muscle, and kidney."

ORIGIN

Alignment Scores:
 Pred. No.: 5.98e-35 Length: 434
 Score: 320.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 96.61% Mismatches: 0
 Query Match: 97.56% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x CA360651 (1-434)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTTPArgAlaLysLysAspAsnHisLeu 20

Db 140 GTGAGGGTTTGACGGCTCAGGCTGTGACCCCTGGCGGCGCAAGAGACACACACTTC 199

QY 21 AsnPhaSnlyAsnAspValIleThrValLeuGluGlnGlnAspMetTyTrpPheGly 40

Db 200 AACTTCAACAGACGACGATGATCAGTCTGAGAGCAGCAGCAATGTGTGTTCGCGC 259

QY 41 GluValGlnGlyGlnLysGlyTyTrpPheProLysSerTyValLysLeuIleSerGly 59

Db 260 GAGGTGAGGGCGCAGCGGCTGTGTTCCCAAGTCTTACGTAAGCTCATCTGTGT 316

RESULT 14

BO179493 690 bp mRNA linear EST 30-APR-2002

LOCUS UT-M-EMO-bww-k-03-0-UT_r1 NIH_BMAP_EMO Mus musculus cDNA clone

DEFINITION IMAGE:5704058 5', mRNA sequence.

ACCESSION BO179493

VERSION BO179493.1 GI:20354985

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 690)

AUTHORS NIH-MGC <http://nigc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs@remail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.lnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..690

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5704058"

/issue_type="whole brain"

/dev_stage="embryo 15.5 dpc"

/lab host="DH10B (T1 phage resistant)"
 /clone lib="NH1_BMAP_EW0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail, is GTGGTGGAA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

/db xref="taxon:9913"
 /tissue_type="pooled"
 /lab host="DH10B"
 /clone lib="MARC 6BOV"
 /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
 Library made with RNA pooled from multiple tissues
 including liver, lung, hypothalamus, pituitary, and
 placenta/endometrium."

ORIGIN

ORIGIN

Alignment Scores:
 Pred. No.: 3,01e-33 Length: 690
 Score: 310.00 Matches: 58
 Percent Similarity: 98.33% Conservative: 1
 Best Local Similarity: 96.67% Mismatches: 0
 Query Match: 94.51% Indels: 1
 DB: 13 Gaps: 0

Alignment Scores:
 Pred. No.: 3,03e-33 Length: 694
 Score: 310.00 Matches: 58
 Percent Similarity: 98.33% Conservative: 1
 Best Local Similarity: 96.67% Mismatches: 0
 Query Match: 94.51% Indels: 1
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BQ179493 (1-690)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTTPArgAlaLysLysAspAsnHisLeu 20
 DB 423 GTGGAGGGGCTTACAGGCAAGCCCTGTATCCGTGGAGAGCCAAAGAAAGACAACCACTTTA 482

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTTPArgAlaLysLysAspAsnHisLeu 20
 DB 120 GTGGAGGGGCTTACAGGCAAGCCCTGTATCCGTGGAGAGCCAAAGAAAGACAACCACTTTA 179

QY 20 uAsnpheAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTrrTPpPheG1 40
 DB 483 AAATTTTAACAAAAGTACGTCAACCGTTCTGGAAACACCAAGACATGTGGTGTGG 542

QY 21 Asn-PheAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTrrTPpPheG1 40
 DB 180 AATGTGAACAAAAGACGTCAACCGTCTGGAGAGCCAAAGACACATGTGGTGTGG 239

QY 40 yGIuValGlnGlyGlnLysGlyTYrPheProLysSerTYrValLysLeuIleSergly 59
 DB 543 AGAAGTTCAGAGGTCAGAGGGTGTTCCTCCCAAGTCTTACGTGAACCTCATTTCAAGG 600

QY 40 yGIuValGlnGlyGlnLysGlyTYrPheProLysSerTYrValLysLeuIleSergly 59
 DB 240 AGAAGTTCAGAGGTCAGAGGGTGTTCCTCCCAAGTCTTACGTGAACCTCATTTCAAGG 297

Search completed: July 1, 2004, 19:48:48
 Job time: 1780.84 secs

RESULT 15
 CB435046 694 bp mRNA linear EST 25-MAR-2003
 LOCUS 611770 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION
 ACCESSION CB435046
 VERSION CB435046.1 GI:29215685
 KEYWORDS
 SOURCE
 ORGANISM Bos taurus (cow)

REFERENCE
 AUTHORS Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 694)
 Smith, T.P.L., Roberts, A.J., Ehternkamp, S.E., Chitko-Mckown, C.G.,
 Wray, J.E. and Keeler, U.W.
 A second set of bovine ESTs from pooled-tissue normalized libraries
 Unpublished (2003)
 COMMENT
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross match v0.990329.
 plate: FQY8036 row: 0 column: 9
 Seq primer: GTAATACGACTCACTATAGG3.
 Location/Qualifiers
 1..694
 /organism="Bos taurus"
 /mol_type="mRNA"

FEATURES
 source

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 1459.86 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_999_1062
Sequence: 1 GEEIAQVATSYATATGPEQLT.....RGKKROIQWFPANYKULSP 64

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cg2.1/USPTO.spool.p/US09720934/runat.30062004.064539.13430/app.query.fasta_1.1386
-DB=GenEmbl -OPMT=fastop -SUFFIX=rge -MINMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=200000000
-USER=US09720934.@CEN.1.1.7509.@runat.30062004.064539.13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
GenEmbl:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_ov:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*
15: em_ba:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_sts:.*
28: em_un:.*

29: em_vi:.*
30: em_hcg_hum:.*
31: em_hcg_inv:.*
32: em_hcg_other:.*
33: em_hcg_mus:.*
34: em_hcg_pin:.*
35: em_hcg_rod:.*
36: em_hcg_mem:.*
37: em_hcg_vrt:.*
38: em_sy:.*
39: em_hcgo_hum:.*
40: em_hcgo_mus:.*
41: em_hcgo_other:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	341	100.0	3231	6	BD205037	BD205037 Isolated
2	341	100.0	3241	9	HSU61166	U61166 Human SH3 d
3	341	100.0	3812	10	AF132672	AF132672 Rattus no
4	341	100.0	4321	9	HSW80584	EX38175 Homo sapi
5	341	100.0	5199	6	BD205033	BD205033 Isolated
6	330	96.8	4025	10	AF127798	AF127798 Rattus no
7	330	96.8	5195	6	BD205035	BD205035 Isolated
8	330	96.8	5381	9	AF114488	AF114488 Homo sapi
9	330	96.8	5458	6	BD205034	BD205034 Isolated
10	330	96.8	6439	9	AF114487	AF114487 Homo sapi
11	328	96.2	1996	9	AF180522	AF180522 Homo sapi
12	327	95.9	5287	9	AF064243	AF064243 Homo sapi
13	327	95.9	7247	9	AF064244	AF064244 Homo sapi
14	325	95.3	1133	10	AF169621	AF169621 Mus muscu
15	325	95.3	3723	10	AF132478	AF132478 Mus muscu
16	325	95.3	5145	10	AF132481	AF132481 Mus muscu
17	319	93.5	4103	5	AF032118	AF032118 Xenopus 1
18	318	93.3	78190	9	AC106760	AC106760 Homo sapi
19	318	93.3	179264	2	AC023112	AC023112 Homo sapi
20	303	88.9	171603	9	AC012629	AC012629 Homo sapi
21	258	75.7	76179	2	AL606725	AL606725 Dario rer
22	258	75.7	104334	5	AL606751	AL606751 Zedratish
23	258	75.7	163197	5	BX005416	BX005416 Zedratish
24	258	75.7	182740	2	AC138436	AC138436 Dario rer
25	258	75.7	225535	2	BX470235	BX470235 Dario rer
26	253	74.2	747	6	AR175268	AR175268 Sequence
27	253	74.2	1391	9	AF001630	AF001630 Homo sapi
28	253	74.2	2004	9	AF038189	AF038189 Homo sapi
29	253	74.2	2873	6	AF175271	AF175271 Sequence
30	253	74.2	3594	10	AF132480	AF132480 Mus muscu
31	253	74.2	4053	9	HSU61167	U61167 Human SH3 d
32	253	74.2	4447	6	AX428899	AX428899 Sequence
33	253	74.2	4557	9	AF182199	AF182199 Homo sapi
34	253	74.2	4977	10	AF132479	AF132479 Mus muscu
35	253	74.2	5804	10	AK122480	AK122480 Method fo
36	253	74.2	5828	6	BD167848	BD167848 Homo sapi
37	253	74.2	5828	9	AF248540	AF248540 Homo sapi
38	253	74.2	5938	9	AB033082	AB033082 Homo sapi
39	253	74.2	6091	9	AF182198	AF182198 Homo sapi
40	251.5	73.8	134408	2	AC139628	AC139628 Takifugu
41	244.5	71.1	136937	2	AC139627	AC139627 Takifugu
42	242.5	71.1	113146	2	AC138440	AC138440 Tetradon
43	242.5	71.1	127051	2	AC114895	AC114895 Tetradon
44	237	69.5	2408	3	AY069517	AY069517 Drosophi1
45	237	69.5	3750	3	AF054612	AF054612 Drosophi1

RESULT 1

ALIGNMENTS

BD205037
LOCUS BD205037 3231 bp DNA linear PAT 17-JUN-2003
DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
ACCESSION BD205037.1 GI:33014807
VERSION JP 2002511267-A/5.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 5 16-APR-2002;
COMMENT CEDARS SINAI HEALTH SYSTEM ET AL
OS Homo sapiens (human)
PN JP 2002511267-A/5
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R KORENBERG, XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

FEATURES
source
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1..3231
FT 1..3231 /organism="Homo sapiens (human)"
ORIGIN
Location/Qualifiers
1..3231 /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Alignment Scores:
Pred. No.: 6.94e-33 Length: 3231
Score: 341.00 Matches: 64
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x BD205037 (1-3231)

QY 1 G1yG1uG1u1leA1aG1nVai1leA1aSeTyThrA1aThRg1yPrOg1uG1nLeuThr 20
Db 1487 GGAAGAGAAATTGGCCAGATTATGCTCATACACCGCCAGCCGACGACTACT 1546

QY 21 LeuA1aPrOg1yG1nLeu1leA1aRg1yS1ySaNPrOg1yG1yT1rP1rG1uG1y 40
Db 1547 CTCGCCCTGGTCACTGATTTTATTCGAAAAAAGAACCCAGTGATGTGGGAAGA 1606

QY 41 G1uLeuG1nA1aRg1yLys1ySaRg1n1leG1yT1rP1rPhePrO1aAaStTyTVal1ys 60
Db 1607 GAGCTGCAAGACAGCTGGGAAAAAGCGCCAGATAGCTGTTCACGCTAATTATGTAAAG 1666

QY 61 Leu1euSeRPro 64
Db 1667 CTTCTAAGCCCT 1678

RESULT 2
HSU61166 3241 bp mRNA linear PRI 23-JUN-1996
LOCUS HSU61166
DEFINITION Human SH3 domain-containing protein SHP17 mRNA, complete cds.
ACCESSION U61166.1 GI:1438932
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Sparks, A.B., Hoffman, N.G., McConnell, S.J., Fowlkes, D.M. and Kay, B.K.
TITLE Cloning of 13 gene targets: systematic isolation of SH3 domain-containing proteins
JOURNAL Nat. Biotechnol. 14 (6), 741-744 (1996)
MEDLINE 98294438
PUBMED 9630982
REFERENCE Pirozzi, G., McConnell, S.J., Uveges, A. and Fowlkes, D.M.
AUTHORS Direct Submission
TITLE Submitted (18-JUN-1996) CYTOGEN Corp., 307 College Road East, Princeton, NJ 08540, USA
JOURNAL
FEATURES
source
1..3241
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
37..1599
/codon_start=1
/product="SH3 domain-containing protein SHP17"
/protein_id="AAC50592.1"
/db_xref="GI:1438933"
/translation="MEAEERLKQKEQERKIELEKQKEAQRQAEKQWLEHVOED
EHQPRKLEHEERKRESYVKQDEKQADQKGRLEHQBEPKPAVQAWST
AEKGP1L1SAQENKVVYRALYPFESRSHDIT10PDIYMWDSOTGEWGLGEL
KKGTFMPANAEEKIPENVEVPAVPKPVDTSDSAAPK1AARETAP1A1AVTS
NMADRSSTPTSTNKEPDTNMAA0PSLT1P5AQQLRQRAFPATATGSSPS
VLQGEVEBGLQALYPMRAKQNDN1FNNDV1YLEQDMMFGEVQCKWPFK
SYVKLISGPIRKSTMSDSSESPASIKRVASPAKPVSGSE1AQV1ASTATGPEQ
LTLAPQDILIRKKNPGWMEGLQANGKQIQ1WFPANVYK1LSPGSKITPPEPK
STALAAVCQVGMVDTYAQNDELAFNKGQ11NVLNKEDPMMKGEVNGVGLPFSNY
VKLT1TMDPSQO"

Alignment Scores:
Pred. No.: 6.96e-33 Length: 3241
Score: 341.00 Matches: 64
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x HSU61166 (1-3241)

QY 1 G1yG1uG1u1leA1aG1nVai1leA1aSeTyThrA1aThRg1yPrOg1uG1nLeuThr 20
Db 1159 GGAAGAGAAATTGGCCAGATTATGCTCATACACCGCCAGCCGACGACTACT 1218

QY 21 LeuA1aPrOg1yG1nLeu1leA1aRg1yS1ySaNPrOg1yG1yT1rP1rG1uG1y 40
Db 1219 CTCGCCCTGGTCACTGATTTTATTCGAAAAAAGAACCCAGTGATGTGGGAAGA 1278

QY 41 G1uLeuG1nA1aRg1yLys1ySaRg1n1leG1yT1rP1rPhePrO1aAaStTyTVal1ys 60
Db 1279 GAGCTGCAAGACAGCTGGGAAAAAGCGCCAGATAGCTGTTCACGCTAATTATGTAAAG 1338

QY 61 Leu1euSeRPro 64
Db 1339 CTTCTAAGCCCT 1350

RESULT 3
AF132672 3812 bp mRNA linear ROD 28-JUN-1999
LOCUS AF132672
DEFINITION Rattus norvegicus EH-domain/SH3-domain containing protein mRNA, complete cds.
ACCESSION AF132672
VERSION AF132672.1 GI:4838525

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 3812)
Okamoto, M., Schoch, S. and Sudhof, T.C.
BRSN/interectin, a protein that contains EH and SH3 domains and
binds to dynamin and SNAP-25. A protein connection between
exocytosis and endocytosis?
J. Biol. Chem. 274 (26), 18446-18454 (1999)
99303609
10373452
2 (bases 1 to 3812)
Okamoto, M., Schoch, S. and Sudhof, T.C.
Direct Submission
Submitted (02-MAR-1999) Center for Basic Neuroscience and HHMI,
Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX
75235, USA

Location/Qualifiers
1..3812
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
41..3481
/note="EHSN1"
/codon_start=1
/product="EH-domain/SH3-domain containing protein"
/protein_id="AAD31026.1"
/db_xref="GI:4838526"
/translation="MAQFPFPGSGSLDIALITYEERAKHDQFOSLKIPISGFTTGOA
RNFPSGSLPQVLAQIWMALDNKQGRMDQVFSIAMKLITKLQGYQPLPVM
KOOPAISSAPAFGIGMAGMPLTAVAPVPMQSLPYVQSGPPLVSAVAPALPM
GAPVIOPLPAFAHAPATLPKSSSFSGSGGSLNTLQKAGSDVSAVAPALPM
OSSRLKROLFNSHDKTMSGLTGPORATLIMSSLPVQALASIMVLSIDQKGLTA
EEFLIMHLDIVAMSGOPLPVPVPEYIPSPRRVSGSGMSVSSSSSSSSSSSSSS
SEDEQVKKLPTTFEDKRNERNERNGLEERKQALLBQKREKRLQLEAEQER
KERROEKRLQLEKQRELEKRELEKRELEKRELEKRELEKRELEKRELEKRELE
NRRQELLTQNNKQEGIVLAKRKTLFELALNDKQLEKQLEKQLEKQLEKQLEKQLE
IESTNKSRELIAEITHLQQLQESQOQMLIPEKQILSDQKQVQNSLHDSLT
LKALAKELARQLEQLEDEVEKETRSKLOEIDVFNOLKELREIHSKQLOKQSI
EERLKQEKQERKSLERKQEKQERKQEKQERKQEKQERKQEKQERKQEKQERKQ
EYDVKKKEAEERAKPEVQKOSRLRHHQSPAPAPAPAPPTTEKPLTISAQSAKY
YRKALYPFESRDEITTPGDIVMVKGWVDSQTEGEMLGEPEKGTGMPAPYA
EKIPENELPTPAKPVDTLTSAPAPKALARTAPLPTVTSSEPTTPNNMADFSTVPS
STNEKPEITDMDTMAQPSLTVPASQLRQSRAPATATGSSPVLQGEKEVEGLQ
AALYEPWRAKQNDHNFNKSVDITVLEQODMMWFGVQKQWPKSYKVLISGPYK
STSDTGPTEAPSLKRVASPAKPAIPGEIADQVIASTATGPEQUTLAPQGLILIR
KNRGGWMEGELQARGKROIGWPAKPAIPGEIADQVIASTATGPEQUTLAPQGLILIR
MDITYKQNDDELAFSKQILINVLKSEDPDMWKGESGQVGLFPSNTYKLTITDMDPSQQ
"

ORIGIN
Alignment Scores:
Pred. No.: 8,29e-33 Length: 3812
Score: 341.00 Matches: 64
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x AF132672 (1-3812)

QY 1 GlyGluGluIleAlaGlnValIleAlaSerTyrThrAlaThrGlyProGlnGlnLeuThr 20
Db 3041 GGAGAGAAATCGCCCGATTCCTTCTACACGCTGCTACGGGTCTCTGAACAGCTCAC 3100
QY 21 LeuAlaProGlyGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 40
Db 3101 CTGGCTCTCGGTCAGCTGATTTGATCCGAAAAAGAACCCCGATGATGATGGGAGGA 3160
QY 41 GluLeuGlnAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 60

Db 3161 GAACTCAAGCTCGAGGAGGAAAGCCGACATGAGGTTCCAGCAAAATATGTCAA 3220
QY 61 LeuLeuSerPro 64
Db 3221 CTTCTAAGCCCT 3232

RESULT 4
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

HSMB06384
Homo sapiens mRNA; cDNA DKFZp686J17173 (from clone DKFZp686J17173).
BX538175.1 GI:31874620
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4321)
Fabo, G., Han, M. and Wiemann, S.
Direct Submission
Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp686J17173) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heuberberg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

Location/Qualifiers
1..4321
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="21q22.11"
/clone="DKFZp686J17173"
/tissue="human uterus endometrium endothel primary cell culture"
/clone_1b="686 (synonym: hlc03). Vector p8port1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
1..4321
/gene="DKFZp686J17173"
<1..1886
/gene="DKFZp686J17173"
/note="intersectin long isoform, N-terminus truncated,
differentially spliced"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAD98050.1"
/db_xref="GI:31874621"
/translation="PFFPFLIVVHLKLAPVIOPLPAFAHAPATLPKSSPSRSG
PQSQNTKQKASFPVAVSYPVPAENAVQSSSLKRLQPLNSDKMTSGTLTPQART
ILMQSLPQQLASITWNLSDIDQDKLTLBEFLIALMLDIVANSQPLPVLPEYIP
PSFRVRSQSGISVISSTVDQRLPEEPLVEDQQLLEKLPTEFDEKRENERGL
ELEKROALIEQKREKRELAQLEKRAEKEKREKREKREKREKREKREKREKREKRE
OREBERKELEKREKREKREKREKREKREKREKREKREKREKREKREKREKREKRE
FELAEKNDKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQ
LEKLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQ
KLOEIDITFNQLEKLEIRHNKQDQKQKSEAEERLQKQEKREKLEKREKREKRA
QERDKQMLHVOQDEHQRPRKHEEKREKREKREKREKREKREKREKREKREKREKRA
HOPPAKPAVQAPWSTAEGKPLTISAQENKVVVYRALYPFESRSHDEITTPGDIWV
DES"
4267..4272
/gene="DKFZp686J17173"
4269
/gene="DKFZp686J17173"

ORIGIN
Alignment Scores:

Pred. No.: 9.5e-33
 Score: 341.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 9
 Gaps: 0

US-09-720-934-2_copy_999_1062 (1-64) x HSM806384 (1-4321)

QY 1 GlytugluilualaglnvalillealaserYtrhratarnrglyProgluglnleuthr 20
 Db 2562 GGAGAGAAATTTGCCAGCTTATTCCTTCATACACGCCGCCGCCAGAGAGCTCACT 2621
 QY 21 LeuilaProglYglnleullelleullearYglsysaenProglYgltYtrTroglnugly 40
 Db 2622 CTCGCCCTGGTCAGCTATTTCATCCGAAAAAGAACCCAGGTGATGTTGGAGAGGA 2681
 QY 41 GluueuglnalAarGgLYlsYlsAarglnlleglyTtrPheProalAasnTyVallys 60
 Db 2682 GAGCTGCAAGACGCGGAGAAAGCCGACATAGCGCTGTTCCCGCTAATTATGTAAG 2741
 QY 61 LeuIeuSerPro 64
 Db 2742 CTTCTAAGCCCT 2753

RESULT 5
 BD205033
 LOCUS BD205033 5199 bp DNA linear PAT 17-JUL-2003
 DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
 ACCESSION BD205033
 VERSION BD205033.1 GI:33014803
 KEYWORDS JP 2002511267-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Korenberg, J.R. and Chen, X.N.
 TITLE Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
 JOURNAL Patent: JP 2002511267-A 1 16-APR-2002;
 COMMENT CEDARS SINAI HEALTH SYSTEM ET AL
 OS Homo sapiens (human)
 PN JP 2002511267-A/1
 PD 16-APR-2002
 PF 16-APR-1999 JP 2000543610
 PR 16-APR-1998 US 60/082007
 PI JULIE R. KORENBERG, XIAO NING, CHEN
 PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
 PC C12Q1/68,
 PC G01N33/68//A61K48/00, C12N15/00, C12N5/50
 CC Isolated SH3 gene relating to myeloproliferative disorders and

FEATURES
 Source
 CC and utilization thereof.
 FH Key Location/Qualifiers
 FT source 1..5199
 FT Location/Qualifiers
 1..5199 /organism="Homo sapiens (human)"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1.16e-32
 Score: 341.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 6
 Gaps: 0

US-09-720-934-2_copy_999_1062 (1-64) x BD205033 (1-5199)

QY 1 GlytugluilualaglnvalillealaserYtrhratarnrglyProgluglnleuthr 20
 Db 3202 GGAGAGAAATTTGCCAGCTTATTCCTTCATACACGCCGCCGCCAGAGAGCTCACT 3261
 QY 21 LeuilaProglYglnleullelleullearYglsysaenProglYgltYtrTroglnugly 40
 Db 3262 CTCGCCCTGGTCAGCTATTTCATCCGAAAAAGAACCCAGGTGATGTTGGAGAGGA 3321
 QY 41 GluueuglnalAarGgLYlsYlsAarglnlleglyTtrPheProalAasnTyVallys 60
 Db 3322 GAGCTGCAAGACGCGGAGAAAGCCGACATAGCGCTGTTCCCGCTAATTATGTAAG 3381
 QY 61 LeuIeuSerPro 64
 Db 3382 CTTCTAAGCCCT 3393

RESULT 6
 AF127798
 LOCUS AF127798 4025 bp mRNA linear ROD 28-JUN-1999
 DEFINITION Rattus norvegicus EH- and SH3-domain containing protein EHSN1 mRNA, complete cds.
 ACCESSION AF127798
 VERSION AF127798.1 GI:4835852
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS Okamoto, M., Schoch, S. and Sudhof, T.C.
 TITLE EHSN1/Intersectin, a protein that contains EH and SH3 domains and binds to dynamin and SNAP-25. A protein connection between exocytosis and endocytosis?
 JOURNAL J. Biol. Chem. 274 (26), 18446-18454 (1999)
 PUBMED 99303609
 REFERENCE
 AUTHORS Okamoto, M., Schoch, S. and Sudhof, T.C.
 TITLE Direct Submission
 JOURNAL Submitted (11-FEB-1999) Center for Basic Neuroscience and HHMI, Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235, USA

FEATURES
 Source
 CDS

Location/Qualifiers
 1..4025 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 41..3694 /codon_start=1
 /product="EH- and SH3-domain containing protein EHSN1"
 /protein_id="AAD30271.1"
 /db_xref="GI:4835853"
 /translation="MAQPTPPGSGSLDIWATVERAKHQDQFSLKPISEFTSGDQ
 RNPFQSLGPPVLAQIOWALADNMKGDMDVBSIAKLILKILQGLQPPALPDV
 KOAPALISAPARFAGIGMAAGMPPLTAVAPVMSGSLPVVMSPLVSSVPAQAPVPLAN
 GAPVLOPLPAPFAPATLPKSSSFSSGSGPSQNTLQKXQSGSTVYAPAPAAEAVP
 QSSRLKRYQLPNSHDKTWSGHLTGPQATILMOSLSLPAQLASLWNLSIDITQDGKTLA
 BEFLIAMHLIDVAMSGQPLPVLPEPEYIPSEFRVRSSGMSVSISSSADQLPEPS
 SEDEQVVEKALPVTPEDEKRNPFERGNILEKRSQALEQCKRQERLAQLERAQER
 KERERQERQERQERLEKQERLEKQERLEKQERLEKQERLEKQERLEKQERLEKQER
 NRROELLTORRKOQEGIVLCKARKKTLFEFLAEALNDKKHLEKGLQDIRCLAQORQ
 IESTNKSRELIARITHTHQQLQSSQQLGSLIPEKQILSLQQLQVQONSISHRSLIT
 LKRALBAKELARQQLREQLDEVERKTRSKQLQETIVFNQLKELKISHKQQLQKRSI
 EAERLKOQEKQERKSLLEKQEKQERQVQERDKQWQHVQEQQRPKPEBQDLKRI
 EDSYKKEAEERAKPEVODKQSLFHPQEPKAKPAQAPWPTTEKGPGLTISAQSAKVY
 YRALYFPESRSHDEITIQPDIVWVGQEWVDSQTEGQWLGEPKGTQWFPANVA
 EKIPENILPMPAKVTDTLSAPAKLALRETPALPVTSSSRSTTPNNWADPSSVWPS
 STNKEPDTNDMTAAQPSLTVPSAQRQSGATTPATPATSSSPKSPVGLQGEQKSGVGLQ
 AQLALYPMRAKDNHLNFRKSDVITVLEQQDMMWPFQVGQKGMFKPSVKLISGVPRK

ORIGIN

STSIDTGPTEAPSSLRKVASPAKPAIPGEFVAMTYESESHEGLJTFQGHVIVTK
KGGDMWGTGTVGSGVPSXVYVLRKDSBEGSTGKXGLEKREINQVIASTYATPE
QUTLAPQQLIRKKNPGWMBSELQKRGKROIGMPANVYVLSBGTISKITPTELP
KTAVQPAVCQVIGMIDYTAQNDELAFSKQIINVLKSDPDWVKGEVSGVGLFPSN
YVLTITDMDPSQQ"

Alignment Scores:

Pred. No.: 2,13e-31 Length: 4025
Score: 330.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.77% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AF127798 (1-4025)

QY 3 GlnAlaGlnValIleAlaSerYrThAlaThGlyProGluGlnLeuThLeuAla 22
Db 3260 GAATGCGCCAGGCTATTGCTTCTCTACACTGCTACGAGCTCTGAAACGCTCAGCTGCT 3319
QY 23 ProGlyGlnLeuLeuLeuIleArgLySAsnProGlyYrTrpTrpGluGlnLeu 42
Db 3320 CCGTGTGAGCTGATTTTGATCCGAAAAAGAACCCCGGTGATGTCGGAAGGAACTA 3379
QY 43 GlnAlaArgGlyLySAsnArgGlnIleGlyTrpPheProAlaAsnTyValLySLeuLeu 62
Db 3380 CAAGCTGAGGGAAGAAAGCCGACAGATGAGGTGTCCGACCAATATGTCAAACTTCTA 3439
QY 63 SerPro 64
Db 3440 AGCCCT 3445

RESULT 7
LOCUS BD205035 5195 bp DNA linear PAT 17-JUN-2003
DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof.
ACCESSION BD205035
VERSION BD205035.1 GI:33014805
KEYWORDS JP 2002511267-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 5195)
AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 3 16-APR-2002;
COMMENT CEDARS SINAI HEALTH SYSTEM ET AL
OS Homo sapiens (human)
PN JP 2002511267-A/3
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JUIIE R KORENBERG, XIAO NING CHEN
PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

FEATURES

source

CC leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1..5195
FT Location/Qualifiers
FT 1..5195 /organism='Homo sapiens (human)'.
1..5195
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN

Alignment Scores:

Pred. No.: 2,81e-31 Length: 5195
Score: 330.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.77% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x BD205035 (1-5195)

QY 3 GlnAlaGlnValIleAlaSerYrThAlaThGlyProGluGlnLeuThLeuAla 22
Db 3452 GAATGCGCCAGGCTATTGCTTCTCTACACTGCTACGAGCTCTGAAACGCTCAGCTGCT 3511
QY 23 ProGlyGlnLeuLeuLeuIleArgLySAsnProGlyYrTrpTrpGluGlnLeu 42
Db 3512 CCGTGTGAGCTGATTTTGATCCGAAAAAGAACCCCGGTGATGTCGGAAGGAACTG 3571
QY 43 GlnAlaArgGlyLySAsnArgGlnIleGlyTrpPheProAlaAsnTyValLySLeuLeu 62
Db 3572 CAAGCAGTGGGGAAGAAAGCCGACAGATGAGGTGTCCGACCTATATGTAAGCTTCTA 3631
QY 63 SerPro 64
Db 3632 AGCCCT 3637

RESULT 8

LOCUS AF114488 5381 bp mRNA linear PRI 16-JUN-2002
DEFINITION Homo sapiens intersecin short isoform (ITSN) mRNA, complete cds.
ACCESSION AF114488
VERSION AF114488.1 GI:4808824

SOURCE

Homo sapiens (human)

REFERENCE

1 (bases 1 to 5381)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Pucharcos, C., Fuentes, J.J., Casas, C., de la Luna, S., Alcantara, S.,
Arbones, M.R., Soriano, E., Estivill, X. and Pritchard, M.
TITLE Multivalent binding protein expressed in proliferating and
differentiating neurons and overexpressed in Down syndrome
JOURNAL Eur. J. Hum. Genet. 7 (6), 704-712 (1999)

JOURNAL MEDLINE 99415290
PUBMED 10482960
REFERENCE 2 (bases 1 to 5381)
AUTHORS Pucharcos, C., Fuentes, J.J., Pritchard, M. and Estivill, X.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
Cancer Research Institute, L'Hopitalat de l'lo., Avia.
Castelldefels km. 2,7, Barcelona 08907, Spain
Location/Qualifiers
1..5381
/organism='Homo sapiens'
/mol_type='mRNA'
/db_xref='taxon:9606'
/chromosome='21'
/map='21q22.1-q22.2'
1..5381
/gene='ITSN'
/gene='ITSN'
269..3931
/gene='ITSN'
/codon_start=1
/product='intersecin short isoform'
/protein_id='AAD29953.1'
/db_xref='GI:4808825'

FEATURES

source

CC leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1..5381
FT Location/Qualifiers
FT 1..5381 /organism='Homo sapiens (human)'.
1..5381
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

EEFTLHMLIDVAMSFOPLPVLPEXYIPSPFRVRSNGSISVTSSTVQORLEPEPV
 LEDEQOQLEKLPVTEDEKRENFERNLEKRRQALLBEQKREGERLADLEAEOE
 RKERERQOERKQLELEKQERLEKQERERERKIERERAKBLEQOLEME
 RNRQOELINDNKEQEDIVLKARKKLELEFELALNDKQLEKLDIDIRLRTQO
 ETEKSTRELARLAEITHLQOLESQOMKLIPEKQILNDQIKOYQNLHSDSLV
 TIKALKEXELARQHLNDQLEVEKERTSKQILIDIPNOLKELEIHNKQILKQKS
 MEARLKQERQERKILEKQKEARQARQOEWLEHVEQEHORPRRLIHEEKV
 KRESVYKKQSEERKQEOADKLGRLPHQHEPAPKPAVOAWSTAREKGLTISAENV
 KVVYRALYPPERSHDEITIQPDIVNKGEMVDESQTEGEMWLGELKRTGMFPA
 NYAEKIPENEPAPVAPVTDSTAPAPKLAREPAPLAVTSSEPTTPNMAFSSST
 WPTSTNEKPEITDNDVMAWAPSLTPSPAGQLRQSAFTPATATSSSPVIGQEKYE
 GLQALYPMFAKQDNLNFKNDVITVLEQDMMWFEVQCGKFPKSVKLIQGP
 IRLKSTMDSGSSESPSLKRVASPAKVVSEEFIAVYTESSEQGLTFPOQGVLL
 VTKDQDMWNTGVDKAGVPSNVWYIKDSESGAGTGTGSGAAGAGAGCTG
 GPEQTLAPQGLILIRKKNPGGMYEGLQAGKRRQIGPFRANYVKKLSPTSTCTTP
 EPPSTALAAVCVIAIMDYATANDDELAFNKQIINVLNKDEPDMWKGENVQVGLF
 PSNVYKLTITMDPSQ"

ORIGIN

Alignment Scores:
 Pred. No.: 2.97e-31 Length: 5381
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: 9 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AF114488 (1-5381)

QY 3 Glut1ealaginVal1leal1aSerTyrThra1aThrGlyProGluGlnLeuThreVala 22
 Db 3497 GAAATTGCCAGGTTATTGCTCATACACCGCACCGCCCCGAGCAGCTCACTCTCGCC 3556
 QY 23 ProGlyGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 42
 Db 3557 CCTGTCAGCTGATTGATTCGAAAAAGAACCCAGGTGATGTTGGAGAGAGAGCTG 3616
 QY 43 Gln1aArGlyLysLysArGln1leGlyTrrPheProAlaSerTyrValLysLeuLeu 62
 Db 3617 CAAGACCTGGGAAAAAGCGCCGATAGGCTGGTTCCTCCAGCTAATTATGTAAGCTTCTA 3676
 QY 63 SerPro 64
 Db 3677 AGCCCT 3682

RESULT 9

BD205034 5458 bp DNA linear PAT 17-JUL-2003
 LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
 DEFINITION leukemia and utilization thereof.

ACCESSION BD205034.1 GI:33014804
 VERSION BD205034.1
 KEYWORDS JP 2002511267-A/2.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5458)
 AUTHORS Korenberg, J.R. and Chen, X.N.
 TITLE Isolated SH3 gene relating to myeloproliferative disorders and
 JOURNAL leukemia and utilization thereof
 PATENT: JP 2002511267-A 2 16-APR-2002;
 CEDARS SINAI HEALTH SYSTEM ET AL

COMMENT OS Homo sapiens (human)
 PN JP 2002511267-A/2
 PD 16-APR-2002
 PF 16-APR-1999 JP 2000543610
 PR 16-APR-1998 US 60/082007
 PI JULIE R. KORENBERG, XIAO NING CHEN
 PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
 PC C12Q1/68,
 PC GOIN33/68//A61K48/00, C12N15/00, C12N5/00
 CC Isolated SH3 gene relating to myeloproliferative disorders and

CC CC and utilization thereof.
 FH FH Key Location/Qualifiers
 FT FT source 1..5458
 /organism="Homo sapiens (human)"

FEATURES
 source 1..5458
 Location/Qualifiers
 1..5458
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 2.97e-31 Length: 5458
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: 6 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x BD205034 (1-5458)

QY 3 Glut1ealaginVal1leal1aSerTyrThra1aThrGlyProGluGlnLeuThreVala 22
 Db 3495 GAAATTGCCAGGTTATTGCTCATACACCGCACCGCCCCGAGCAGCTCACTCTCGCC 3554
 QY 23 ProGlyGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 42
 Db 3555 CCTGTCAGCTGATTGATTCGAAAAAGAACCCAGGTGATGTTGGAGAGAGAGCTG 3614
 QY 43 Gln1aArGlyLysLysArGln1leGlyTrrPheProAlaSerTyrValLysLeuLeu 62
 Db 3615 CAAGACCTGGGAAAAAGCGCCGATAGGCTGGTTCCTCCAGCTAATTATGTAAGCTTCTA 3674
 QY 63 SerPro 64
 Db 3675 AGCCCT 3680

RESULT 10

AF114487 6439 bp mRNA linear PRI 16-JUL-2002
 LOCUS Homo sapiens intersecctin long isoform (ITSN) mRNA, complete cds.
 DEFINITION AF114487
 ACCESSION AF114487
 VERSION AF114487.1 GI:4808822

KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 6439)

REFERENCE 1 (bases 1 to 6439)
 AUTHORS Pucharcos, C., Fuentes, J.J., Pritchard, M. and Estivill, X.
 TITLE Direct Submission
 JOURNAL Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
 Cancer Research Institute, L'Hospitalet de llo., Avia.
 Castelldefels km. 2.7, Barcelona 08907, Spain

FEATURES
 source 1..6439
 Location/Qualifiers
 1..6439
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.1-q22.2"
 1..6439

gene

```
/gene="ITSN"
269..5434
/gene="ITSN"
/codon_start=1
/product="intersecin long isoform"
/protein_id="A029952.1"
/db_xref="GI:4808823"
/translation="MAQPPPPGSGSLDINALTYEBEAKNDQOFHSIKPISGFTTGOA
RNFQSGLEPQVLAQIMALADKNDGNDQVFSIAKLIKLGQYGLPSPVPM
KQPVVAISSAPAFMGSLAMPPLTAVAPVPMGSIIVVMGSPVLVSVPAAPIPLAN
GAPVIOPLPAFAPPAATLPKSSSFSGSGSLQNLTKQAGSFDVASVPVMAVAP
OSSRLKYROLFNSHDKTMSGHLTGPOARTILMOSSLPQALATNMLSIDODGKTLA
EERILAMHLIDVMSGOLPEVLPPEVIPSPPRVSISGISTYSTDQKLEPPEV
LEBQOQLEKLEPVTEDKKRENFERNLEERRALLBQOKSEORLAQORAE
KERERQOEKROLELEKQERLEKREKREKLEERRAARELEKROKLEWE
RNRORLELNQNRKOEIDIVLKAKKTLERLEALNDKQLEKQODIRCRILTRQ
EISTNSRELRJAETHTLOOLOSQOMGRILPKOILNDQLKQVQONSRLHSDLV
TLKRALEAKELAROHLDOLDEVEKERSRLOEIDIPNNOLKELEIHNKQOLKREKL
MEARLKQKOEORKEITIEKQKEAORRAOERPKMLEHYQOEDEORPRKLEBEKL
KRESVYKDKGEKQOEADKGRLEFHOEOPAKFAVQAPMSTAEKGPLTISAQENV
KVVYVALYPESRSHDEITTOGDIVMWGEWVDSQTEGPMGLGELKGTGMPA
NVAEKIPENEVPAPVPVDTSTAPAPKALRTAPPLAVTSESPSTPNNDWAFEST
WPTSTNEKETNDWMAAOPSLTVPASQORSAFTPATATSSSPPLVAGGEVYE
GLOAOLYMPBAKNDHNLNPKNDVITVLEQOMMMWGEVQOGKMPKSVYKLISG
IRKSTMSGSSSPASLKRVAAPKPVSGEFTAMTYESSBOGDLTPOCGDVL
VIRKSDMMWTGYGDKAGVPPSNVYKLKDSBSGCTGKTSUGKRELIQVASTAT
GEOULTAEQGLILIRKKNPGWMBELQARGKROIGWFPANVYKLISGRTIPT
EPKSTALAAVCQVIMDYTAQNDDELAFNKQIINVINKEDPMWKBEVMOQVLEF
PSNVYKLITDMDPSQOMCSDLHLDMLPTERKQGYIHLLIYENYVNDQLVTEI
FOKLESELTKEKEVAMI FVNMKELIMCINKILKALRVKMKSGEMVMDJLIS
AOLPHMOPYIRECSROLNGANALIOCKTDEAPDKERFKLAMPKCKMPLSFTLK
MORATRYPLIKILKILENTPBNDHSHLHAKELABELCSQNEGVREKENSRLMT
QAVOCEGSEQLVFNSTVNCJGPRKPLFHSGLKYKAKSNELGFLFNFPLLTQTLK
PLSSGSTDVFSKSNLYQMYKTPILFNEVLYKLPDPSGDEPIFIHSHIDVYLR
AASINERTAMVCKIKAASELYLETEKKEKALVRSORATIGRLMVAWVEGIELEK
CSHGKSNPYCEVTWMSOCHITKTIDOTNPKNNSCOPIRDLBOEVCIITVFEEDQ
FSPDPLGSTEIRVADIKKQDSKGPVTKCLLHEVPTGEIVRDLQLDEDE"
```

ORIGIN

Alignment Scores:

Pred. No.:	3..55e-31	Length:	6439
Score:	330.00	Matches:	62
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	96.77%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x AF114487 (1-6439)

```
QY      3  GlnuLEAlaGlnuValleAlaSerTYrThraLathGlyProGluGlnleuThrLeuA 22
          |||||
Db      3497  GAAATTGCCAGGTTATTGGCTCATACACGCCACCGGCCCGCAGAGCTCCTCGCC 3556

QY      23  ProGluGlnleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 42
          |||||
Db      3557  CCGGTCAGCTGATTTTGGATCCGAAAAAGAACCCAGGTGATGCGGGAAGAGACTG 3616

QY      43  GlnAlaArgGlyLysLysArgGlnIleGlyTYrPheProAlaAsnTYrValLysLeu 62
          |||||
Db      3617  CAAGCAGCGTGGAAAAAGCGCAGATAGGCTGATTCACACTAATTATGTAAGCTTCTA 3676

QY      63  SerPro 64
          |||||
Db      3677  AGCCCT 3682
```

RESULT 11

AF180522

LOCUS AF180522 1996 bp mRNA linear PRI 05-SEP-1999

DEFINITION Homo sapiens intersecin short form 2 (ITSN) mRNA, partial cds.

ACCESSION AF180522

VERSION AF180522.1 GI:5823551

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1996)

AUTHORS Tsyba,L.O., Kvasha,S.M., Skripkina,I.Y., Anoprienko,O.V., Slavov,D., Tassone,F., Rynditch,A.V. and Gardiner,X.

TITLE Mouse homolog of human chromosome 21 genes

JOURNAL Unpublished

2 (bases 1 to 1996)

REFERENCE Tsyba,L.O., Kvasha,S.M., Skripkina,I.Y., Anoprienko,O.V., Slavov,D., Tassone,F., Rynditch,A.V. and Gardiner,X.

AUTHORS Direct Submission

TITLE Submitted (25-AUG-1999) Department of Molecular Oncogenetics, Institute of Molecular Biology and Genetics of National Academy of Sciences of Ukraine, 150, Zabolotnogo str., Kiev 252627, Ukraine

JOURNAL Location/Qualifiers

1..1996

FEATURES

source

gene

CDS

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/cisue_type="brain"
/dev_stage="infant"
<1..1996
/gene="ITSN"
<1..566
/codon_start=3
/product="intersecin short form 2"
/db_xref="GI:5823552"
/translation="KSYVKLISGPIRKSTMSGSSSPASLKRVAAPKPVSGEE
IQAVISYTAAGSEQLTAPGQILIRKKNPGRWMBELQARGKROIGWFPANVYKL
LSGRTKIPTEPKSTALAAVCQVIMDYTAQNDDELAFNKQIINVINKEDPMWK
KBEVMOQVGLFSPSNVYKLITDMDPSQO"
```

ORIGIN

Alignment Scores:

Pred. No.:	1..79e-31	Length:	1996
Score:	328.00	Matches:	62
Percent Similarity:	96.88%	Conservative:	0
Best Local Similarity:	96.88%	Mismatches:	2
Query Match:	96.19%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x AF180522 (1-1996)

```
QY      1  GlyuGluuLeuLeuAlaGlnuValleAlaSerTYrThraLathGlyProGluGlnleuThr 20
          |||||
Db      126  GGAAGAAGAAATTGCCAGGTTATTGCTCATACACGCCCGGCCCGCAGAGCTTACT 185

QY      21  LeuAlaProGluGlnleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
          |||||
Db      186  CTCGCCCTGCTGATGATTTTGGATCCGAAAAAGAACCCAGTGTGATGCGAAGCA 245

QY      41  GlnLeuGlnAlaArgGlyLysLysArgGlnIleGlyTYrPheProAlaAsnTYrValLys 60
          |||||
Db      246  GAGCTCAAGCAGCGTGGAAAAAGCGCAGATAGGCTGATTCACACTAATTATGTAAG 305

QY      61  LeuLeuSerPro 64
          |||||
Db      306  CTTCTAAGCCCT 317
```

RESULT 12

AF064243

LOCUS AF064243 5287 bp mRNA linear PRI 21-NOV-1998

DEFINITION Homo sapiens intersecin short form mRNA, complete cds.

ACCESSION AF064243

VERSION AF064243.1 GI:3859852

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5287)
AUTHORS Guipponi, M., Scott, H.S., Antonarakis, S.E.
TITLE Two isoforms of a human brain-specific alternative splicing in a stop codon
JOURNAL Genomics 53 (3), 369-376 (1998)
MEDLINE 99017974
PubMed 9799604
REFERENCE 2 (bases 1 to 5287)
AUTHORS Guipponi, M., Scott, H.S., Antonarakis, S.E.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue Michel-Servet, Geneva 4 CH-1211, Switzerland
FEATURES
source 1..5287
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="fetus"
107..3769
/codon_start=1
/product="intersectin short form"
/protein_id="AAC78610.1"
/db_xref="GI:3859853"
/translation="MAQPTPFGSLDIWALTVEERAKHDOQFSLKPSIGFTIGDOA
RNFPGSLGPOPVLAIQWALADNMNDMDVRSIMKLIKIKLOGVQLSALPVM
KOOPVALISSAPPFGMGISAPPLTAVAPVMSGIPVGMSPITVSSVPTAAVPLAN
GAPVYIOPUPAFAPRATLPKSSSFSSSGPSQNTYKQKQSDIVASVPPVAAWAV
QSSRLKRYQLNSHDKTWSGHLTPQATTIMQSSSLPQAQIASIWNISDIDQDKLTA
BEFLIAMLIDVANSQGLPVLPPPEYIPPSFRVRSQSGISVISTSVDRLEBPV
LEDEOQOLEKLPVTFEDKKRKNFERGNLEKEROALIEQKROERLAQLEBAOE
RKEROEROERKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEK
RNEROLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEK
EIESTNSRELRIAEITHLQOQLOESQOMGRILPEKQILNDQIKOYQNSLHSDIV
TLKALAEKELARQHLRDQDEVEKERSKIQEIDIPNNQKEIRELHNKQLOKQKS
MEARLKQKQERKILIELEKQERAOERAOERAOERAOERAOERAOERAOERAOER
KREESVKKQGEKQ
KVYVYRALYPPESRSHDEITOPGGIIVWKEWDESGOTGPGMLGKGLKGTQMEPA
NYAKIPEWPAVYKPYTDSAPAPKALRETPAPALRETPAPALRETPAPALRETPAP
WPTSTNEKPEPTDMDMAAQSILTVPSAGQLRQSAFIPATATSSSPVIGQEKYS
GLQAOALYPMRAKQ
IRKSTMSQSGSSSPASIKRVAAPAKVVSGBEFIMATYSSSEQGLTPQOQDVL
VTIKSDGMMTGTVDGKAGVPSNRYRLKDSGSGTAGTKGSLGKKEIAQVIASTYAT
GPBOLTLAPQGLILIRKKNPGWMEGLQARKKRQIGMPFANYVYKILNPTSTITPT
EPKSTALANCVIIGMDYTAQNDDELAPRKQILINVLNKEDPDWKKGVNGQVGLP
PSNYVYKLTTHDMSQ"

misc_feature 167..406
/note="encodes BH domain"
misc_feature 767..1936
/note="encodes BH domain"
misc_feature 2324..2524
/note="encodes SH3 domain"
misc_feature 2843..3019
/note="encodes SH3 domain"
misc_feature 3110..3286
/note="encodes SH3 domain"
misc_feature 3326..3520
/note="encodes SH3 domain"
misc_feature 3569..3748
/note="encodes SH3 domain"

ORIGIN
Alignment Scores:
Pred. No.: 6.84e-31
Score: 327.00
Percent Similarity: 100.00%
Best Local Similarity: 98.39%

Length: 5287
Matches: 61
Conservative: 1
Mismatch: 0

Query Match: 95.89% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AF064243 (1-5287)

QY 3 Glut1leaglVal1leal1aSerTyThralaThnglyProgluInleuThreula 22
DB 3335 GAATTCGCCAGGTTATTCCTCATACCCGCCAGCCGCCAGCAGCTACTCTGCC 3394
QY 23 ProgluInleu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu 42
DB 3395 CCTGGTCAGCTGATTTGATTCGAAAAAGAACCAAGCGTGATGCGGAGAGAGCTG 3454
QY 43 Glu1laArgGlyLy1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu 62
DB 3455 CAAGACCTGGGAAAAAGCCAGTAGCGTGTTCCAGCTAATTATGTAAGCTTCA 3514
QY 63 SerPro 64
DB 3515 AACCT 3520

RESULT 13
AF064244 7247 bp mRNA linear PRI 21-NOV-1998
LOCUS AF064244
DEFINITION Homo sapiens intersectin long form mRNA, complete cds.
ACCESSION AF064244
VERSION AF064244.1 GI:3859854
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 7247)
Guipponi, M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and
Antonarakis, S.E.
TITLE Two isoforms of a human intersectin (ITSN) protein are produced by
JOURNAL brain-specific alternative splicing in a stop codon
Genomics 53 (3), 369-376 (1998)
MEDLINE 99017974
PubMed 9799604

REFERENCE 2 (bases 1 to 7247)
AUTHORS Guipponi, M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and
Antonarakis, S.E.

TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue
AUTHORS Michel-Servet, Geneva 4 CH-1211, Switzerland
FEATURES
source 1..7247
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="fetus"
107..5272
/codon_start=1
/product="intersectin long form"
/protein_id="AAC78611.1"
/db_xref="GI:3859855"
/translation="MAQPTPFGSLDIWALTVEERAKHDOQFSLKPSIGFTIGDOA
RNFPGSLGPOPVLAIQWALADNMNDMDVRSIMKLIKIKLOGVQLSALPVM
KOOPVALISSAPPFGMGISAPPLTAVAPVMSGIPVGMSPITVSSVPTAAVPLAN
GAPVYIOPUPAFAPRATLPKSSSFSSSGPSQNTYKQKQSDIVASVPPVAAWAV
QSSRLKRYQLNSHDKTWSGHLTPQATTIMQSSSLPQAQIASIWNISDIDQDKLTA
BEFLIAMLIDVANSQGLPVLPPPEYIPPSFRVRSQSGISVISTSVDRLEBPV
LEDEOQOLEKLPVTFEDKKRKNFERGNLEKEROALIEQKROERLAQLEBAOE
RKEROEROERKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEK
RNEROLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEK
EIESTNSRELRIAEITHLQOQLOESQOMGRILPEKQILNDQIKOYQNSLHSDIV
TLKALAEKELARQHLRDQDEVEKERSKIQEIDIPNNQKEIRELHNKQLOKQKS
MEARLKQKQERKILIELEKQERAOERAOERAOERAOERAOERAOERAOERAOER
KREESVKKQGEKQ
KVYVYRALYPPESRSHDEITOPGGIIVWKEWDESGOTGPGMLGKGLKGTQMEPA
NYAKIPEWPAVYKPYTDSAPAPKALRETPAPALRETPAPALRETPAPALRETPAP
WPTSTNEKPEPTDMDMAAQSILTVPSAGQLRQSAFIPATATSSSPVIGQEKYS
GLQAOALYPMRAKQ
IRKSTMSQSGSSSPASIKRVAAPAKVVSGBEFIMATYSSSEQGLTPQOQDVL
VTIKSDGMMTGTVDGKAGVPSNRYRLKDSGSGTAGTKGSLGKKEIAQVIASTYAT
GPBOLTLAPQGLILIRKKNPGWMEGLQARKKRQIGMPFANYVYKILNPTSTITPT
EPKSTALANCVIIGMDYTAQNDDELAPRKQILINVLNKEDPDWKKGVNGQVGLP
PSNYVYKLTTHDMSQ"

CDS

misc_feature 167..406
/note="encodes BH domain"
misc_feature 767..1936
/note="encodes BH domain"
misc_feature 2324..2524
/note="encodes SH3 domain"
misc_feature 2843..3019
/note="encodes SH3 domain"
misc_feature 3110..3286
/note="encodes SH3 domain"
misc_feature 3326..3520
/note="encodes SH3 domain"
misc_feature 3569..3748
/note="encodes SH3 domain"

```

KVVYRYALYPEFSRSHDEITIOEGDIWMKGEWVDSOTGEPGLGGLGKGTGMPDA
NYAEKIPENVEVPAVPKVTUDSTSAIPAKLALREPPALATSSSEPTTPNMADFSST
WPTSTNKPEPTDWDADWAAPSLTIVEAGLRQRSATPTATATSSSPVLCGEKXE
GLQAOALYPRRAKNDNLNTRKNDVITVLBQDMMFGEVQGMFKSVKLISSP
IRKSTMSDSSSSPASKLVASPAAPVVSGBEFAMVYBESQDGLFFQOQDVL
VTKKDGMWTVGDKAGVFPSPNYVRKJDSGGSLAGTSLGKPEPIALVASYAT
GPEOLTAIPGOLIRKKNPGMGWEGELORAGKROIGPAPANYVKILNKGITPT
EPPSTALAAVCVIGMYDYTAONDELATNKGQILNVLKNDPDMWKGWVNOYGF
PSNVKLTITMDPSQWCSLDLMDLTPTRKQGITHELIYTERYVNDLQVTEI
FOYPMSESLITEREVAMITVNMKELITMCIKLIKALVRKKNSGEMFVKMTDILS
AQLPHMQPIYRFSRQNGAALIQOKTDEAPDEKFEVRLMEPRCKGMLSPFILP
MOVTRYPLIKNILENTPENHDSHSLKALKAELCSQVNEGYEKNRSLMEI
QAHVQCGSLSEOLVFNSTNCLGPKRFLHSGKLYKAKNKELYGLPNDLLTQITK
PLSSGTDKXVSPKSNLOVMYKTPPIELNVLYKLPDPBSGDEPIPHISIDRYTIR
AESINERTAVQKIKASSEIYITEKKRKRAIVSQRATIGRLIMVNESEIYLR
CRSHGKSNPYCEVTMSQCHITITLQDTLLPNKNSNOCFFIRLREVEVLCITVFERQ
FSPDPLFGRTEIRVADIKKQGSKGPVKTLHLHEVPTGEIVRLDLQLEDFE"
misc_feature
167..406
/note="encodes EH domain"
misc_feature
767..1936
/note="encodes EH domain"
misc_feature
2324..2524
/note="encodes SH3 domain"
misc_feature
2843..3019
/note="encodes SH3 domain"
misc_feature
3110..3286
/note="encodes SH3 domain"
misc_feature
3326..3520
/note="encodes SH3 domain"
misc_feature
3569..3748
/note="encodes SH3 domain"
misc_feature
3836..4390
/note="encodes GDS domain"
misc_feature
4649..4819
/note="encodes PH domain"
misc_feature
4895..5143
/note="encodes C2 domain"
ORIGIN
Alignment Scores:
Pred. No.: 9.61e-31 Length: 7247
Score: 327.00 Matches: 61
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.39% Mismatches: 0
Query Match: 95.89% Indels: 0
Gaps: 0
DB: 9
US-09-720-934-2_COPY_999_1062 (1-64) x AF064244 (1-7247)
QY 3 GlnlAlaGlnValIleAlaSerYrThrAlaThrglyProgluGlnLeuThrLeuAla 22
Db 3335 GAATATGCCAGGTATATGCTCATACACCGCCACCGCCCGCAGAGCTCCTCGCC 3394
QY 23 ProgluGlnLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42
Db 3395 CCGGTCACTGATTTTGAATCCGAAAAAGAACCCAGGTGATGCTGGAGAGAGACTG 3454
QY 43 GlnAlaArgGlyLysLysArgGlnIleGlyTTPheProAlaAsnTyValLysLeuLeu 62
Db 3455 CAAGCACGTGGGAAAAAGCGCCAGATAGGCTGCTCCAGCTAATATGTAAGCTTCTA 3514
QY 63 SerPro 64
Db 3515 AACCT 3520
RESULT 14
LOCUS AF169621 1133 bp mRNA linear ROD 12-AUG-1999
DEFINITION Mus musculus Ebel protein mRNA, partial cds.
ACCESSION AF169621
VERSION AF169621.1 GI:5731280
KEYWORDS
SOURCE Mus musculus (house mouse)

```

```

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1133)
Tayba,L.O., Kvasha,S.M., Skripkina,I.Y., Anoprienko,O.V.,
Slavov,D., Tassone,F., Rynditch,A.V. and Gardiner,K.
TITLE
Mouse homologues of human chromosome 21 genes
JOURNAL
Unpublished
2 (bases 1 to 1133)
Tayba,L.O., Kvasha,S.M., Skripkina,I.Y., Anoprienko,O.V.,
Slavov,D., Tassone,F., Rynditch,A.V. and Gardiner,K.
TITLE
Direct Submission
JOURNAL
Submitted (15-JUL-1999) Department of Molecular Oncogenetics,
Institute of Molecular Biology and Genetics of National Academy of
Sciences of Ukraine, 150, Zabolotnogo str., Kiev 252627, Ukraine
FEATURES
source
1..1133
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/dev_stage="15-day embryo"
<1..749
/note="similar to Homo sapiens Intersectin"
/codon_start=3
/product="Ebel protein"
/protein_id="A04848.1"
/db_xref="GI:5731281"
/translation="VAKSTSIDTPESPASLKRVASPAKPAIPGEFAMVYBES
EGDLPFOQDGVVTRKDGDMWTVGDKAGVFPSPNYVRKJDSGGSLAGTSLGKPEPIALVASYAT
GPEOLTAIPGOLIRKKNPGMGWEGELORAGKROIGPAPANYVKILNKGITPT
EPPSTALAAVCVIGMYDYTAONDELATNKGQILNVLKNDPDMWKGWVNOYGF
FSPDPLFGRTEIRVADIKKQGSKGPVKTLHLHEVPTGEIVRLDLQLEDFE"
DMWKGVSQVGLFSPSNVYKLTITMDPSQW"
ORIGIN
Alignment Scores:
Pred. No.: 2.31e-31 Length: 1133
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
Gaps: 0
DB: 10
US-09-720-934-2_COPY_999_1062 (1-64) x AF169621 (1-1133)
QY 3 GlnlAlaGlnValIleAlaSerYrThrAlaThrglyProgluGlnLeuThrLeuAla 22
Db 315 GAATATGCCAGGTATATGCTCATACCGCTCCTACGCTGATGCTCCGAAACACTCACCCTGCT 374
QY 23 ProgluGlnLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42
Db 375 CCGGTCACTGATTTTGAATCCGAAAAAGAACCCAGGTGATGCTGGAGAGAGACTG 434
QY 43 GlnAlaArgGlyLysLysArgGlnIleGlyTTPheProAlaAsnTyValLysLeuLeu 62
Db 435 CAAGCTCGAGGAAAAAGCGCCAGATAGGCTGCTCCAGCAATATGTCACAACTTCTA 494
QY 63 SerPro 64
Db 495 AGCCCC 500
RESULT 15
LOCUS AF132478 3723 bp mRNA linear ROD 09-MAR-1999
DEFINITION Mus musculus Ebel protein mRNA, complete cds.
ACCESSION AF132478
VERSION AF132478.1 GI:4378884
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 3723)
Sengar,A.S., Wang,W., Bishay,J., Cohen,S. and Egan,S.E.
AUTHORS

```

TITLE

The EH and SH3 domain Eps proteins regulate endocytosis by linking to dynamin and Eps15

Job time : 1469.86 secs

JOURNAL

EMBO J. 18 (5), 1159-1171 (1999)

MEDLINE

99164083

PUBMED

10064583

REFERENCE

2 (bases 1 to 3723)

AUTHORS

Sengar, A. S., Wang, W., Cohen, S., Bishay, J. and Egan, S. E.

TITLE

Submitted (02-MAR-1999)

JOURNAL

Submitted (02-MAR-1999)

FEATURES

Submitted (02-MAR-1999)

SOURCE

Submitted (02-MAR-1999)

CDS

Submitted (02-MAR-1999)

FEATURES

Submitted (02-MAR-1999)

SOURCE

Submitted (02-MAR-1999)

CDS

Submitted (02-MAR-1999)

FEATURES

Submitted (02-MAR-1999)

SOURCE

Submitted (02-MAR-1999)

CDS

Submitted (02-MAR-1999)

FEATURES

Submitted (02-MAR-1999)

SOURCE

Submitted (02-MAR-1999)

CDS

Submitted (02-MAR-1999)

FEATURES

Submitted (02-MAR-1999)

SOURCE

Submitted (02-MAR-1999)

CDS

Submitted (02-MAR-1999)

FEATURES

Submitted (02-MAR-1999)

SOURCE

Submitted (02-MAR-1999)

CDS

Submitted (02-MAR-1999)

FEATURES

Submitted (02-MAR-1999)

SOURCE

Submitted (02-MAR-1999)

CDS

Submitted (02-MAR-1999)

FEATURES

Submitted (02-MAR-1999)

SOURCE

Submitted (02-MAR-1999)

CDS

Submitted (02-MAR-1999)

FEATURES

Submitted (02-MAR-1999)

SOURCE

Submitted (02-MAR-1999)

1. .3723
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1. .3642
/codon_start=1
/product="Eps1 protein"
/protein_id="A019746.1"
/db_xref="GI:4378885"
/translation="MAQPTFGSLDVAITVEBRAXHQFSLKPIAGFTTGDQ
RNFPEFQSGLPVLAQIWMALMNDGMDQVESIAMKILTKIQGYLPSTLPVM
KQCPVAISSAPFAGIAGISMPPLTAAPVPMGSIPIVGMSPPLVSVPAAPVPLAN
GAPVIOPLPAPAPAPATWPKSSFSRSGSOLNTKLOKASFDVAGAPAPAAWAP
OSSRLKXQPLFNSHDKTMSGHLTGPOARTIMOSSLPQALASIMNSIDIDQDKLTA
EEFLAMHLIVAMSGOPLPVLPPEYIPPSFRFRSSGMSVSSSVDDRLPEPS
SEDEQOPEPKLPVTFEDKRENFERGVLEKROQLBQOKQERLAQLERAOER
KERROEOEAROLELEKORLEKREERERKEIRREAKRELERQOLEWER
NRROELNQRNKEOBGVVLRKARKELEFELEALNDKGLEGLKODIRCLATOROE
IESTNKSRELRLAETHLOOQLOESQMLGRLEPKOILSDLOKQONSJLHPSILIT
LKRLEAKELARCOLRBDLEVERETRSKLEIDVFNNOILKEIRHISKKOOLOKORSL
EALUKOQERKSLLEKQKDAORRVOEDKQMLHNVQOEDPRPRKREDEBTRK
EDSVRKKEERAKEMQDKOSRLPHQEPKATQAPWSTTEKPLTISAQSVK
VYRALYPERSHDEITIOGDIWMDESQTEBGMGLGELKGTGMFPANVAKIP
ENEVPTPAKPYTDLTSAAPKALALETPAPLPVTSSEPTTPNNWADFSSTWSSNE
KPEEDNDMTMAOPSLTVPSAGOLRORAPATATGSSPSPLVQOGEVGLQOAL
YPMRAKKNHINFNKSDVITVLEODQMMWFGVOGKGMFPKSYKILSGPVKSTSI
DTGTESPASUKRVASPAKRAIRBEERFIAMTYRPSROGDLTPOGGDVIVTKKDD
WMTGVGDKSGVFPENYVRLKDSGSGTAGTSGISGKKEBIAQVIASTAAATGPEQTL
APGQILIRKKNPGQWMEGLQARGKRQIGMFPANVYKLSPTGSKITPTELPETAV
QPAVCQVIGMYDYTMQNDDELAFSGQIINVLNKEDPMMKGEVSGQVGLFPPSNYKL
TTMDPSSQ"

ORIGIN

Alignment Scores:

Pred. No.: 8,366-31 Length: 3723
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AF132478 (1-3723)

QY 3 GlnlealaglnvalilallasegtyrthrAlatnrglyProgluglnleuthrleuAla 22

Db 3208 GAAATTGCCAGGTATGCTTCTTACGCTGCTACTGCTCCGAACTCACTCCTGGCT 3267

QY 23 Progluglnleuileuileuileuileuileuileuileuileuileuileuileu 42

Db 3268 CCTGGGACGCTGATCTGATCGGAAAAAGAACCCAGGTGATGGGGAAGAGACTG 3327

QY 43 GlnAlaarggly 62

Db 3328 CAAGCTCAGGAGAAAGCGCAGATAGGCTGTTCCAGCAATTAATGTCAACTTCTA 3387

QY 63 SerPro 64

Db 3388 AGCCCC 3393

Search completed: July 1, 2004, 16:15:23

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 44.815 Seconds
(without alignments)
792.522 Million cell updates/sec

Title: US-09-720-934-2_COPY_999_1062
Sequence: 341
1 GERIAVIVSYATATGPEQLT.....RKKRKGWFPANYVVLSP 64

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool_p/US09720934/runat_30062004_064540_13455/app_query.fasta_1.1386
-DB=Issued Patents NA -QFMT=fastaf -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720934.@CEN_1_1.284.@runat_30062004_064540_13455 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	253	74.2	747	4	US-08-630-915A-39
2	253	74.2	2873	4	US-08-630-915A-193
3	113.5	33.3	1392	1	US-08-475-894-3
4	113.5	33.3	1392	1	US-08-484-710-3
5	113.5	33.3	1392	2	US-08-484-709-3
6	113.5	33.3	1392	2	US-08-474-697-3
7	113.5	33.3	1659	1	US-08-475-894-1
8	113.5	33.3	1659	1	US-08-484-710-1
9	113.5	33.3	1659	2	US-08-484-709-1
10	113.5	33.3	1659	3	US-08-474-697-1
11	113.5	33.3	3143	4	US-08-671-354-1
12	111	32.6	2757	1	US-08-306-691B-48

13	111	32.6	2757	5	PCT-US93-06251-79	Sequence 79, Appl
14	109	31.0	2793	1	US-07-646-537B-1	Sequence 1, Appl1
15	107.5	32.5	1414	4	US-09-023-655-1356	Sequence 1356, Ap
16	106.5	31.2	1457	4	US-08-630-915A-197	Sequence 197, App
17	106.5	31.2	1457	3	US-09-444-053-3	Sequence 3, Appl1
18	104	30.5	3111	4	US-10-014-882-1	Sequence 1, Appl1
19	104	30.5	3518	4	US-10-014-882-3	Sequence 3, Appl1
20	102.5	30.1	767	4	US-09-328-475C-155	Sequence 155, App
21	102.5	30.1	1045	1	US-08-475-894-5	Sequence 5, Appl1
22	102.5	30.1	1045	2	US-08-484-710-5	Sequence 5, Appl1
23	102.5	30.1	1045	2	US-08-484-709-5	Sequence 5, Appl1
24	102.5	30.1	1045	3	US-08-474-697-5	Sequence 5, Appl1
25	102.5	30.1	1895	4	US-09-620-312D-397	Sequence 397, App
26	102.5	30.1	3870	1	US-08-138-641-1	Sequence 1, Appl1
27	102.5	30.1	3870	1	US-08-138-133-1	Sequence 1, Appl1
28	102.5	30.1	3893	1	US-08-138-641-3	Sequence 3, Appl1
29	102.5	30.1	3893	1	US-08-138-133-3	Sequence 3, Appl1
30	100.5	29.5	4079	4	US-09-016-434-1247	Sequence 1247, Ap
31	99.5	29.2	970	4	US-09-833-381-1294	Sequence 1294, Ap
32	98	28.7	2547	4	US-09-765-298A-9	Sequence 9, Appl1
33	98	28.7	4079	4	US-09-016-434-1246	Sequence 1246, Ap
34	98	28.7	4242	4	US-09-023-655-1201	Sequence 1201, Ap
35	96.5	28.3	822	4	US-09-328-475C-187	Sequence 187, App
36	96.5	28.3	1024	4	US-09-328-475C-3	Sequence 3, Appl1
37	95.5	28.0	1458	2	US-08-942-423-68	Sequence 68, Appl
38	95.5	28.0	2003	2	US-08-942-423-1	Sequence 1, Appl1
39	95.5	28.0	2003	4	US-08-630-915A-25	Sequence 25, Appl
40	95	27.9	5398	3	US-09-356-952-11	Sequence 11, Appl
41	94	27.6	4762	4	US-09-300-958A-30	Sequence 30, Appl
42	93	27.3	467	4	US-09-300-958A-28	Sequence 28, Appl
43	93	27.3	1382	4	US-09-820-005-1	Sequence 1, Appl1
44	92	27.0	1349	4	US-09-023-655-1176	Sequence 1176, Ap
45	91.5	26.8	3345	4	US-08-630-915A-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-08-630-915A-39
; Sequence 39, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLES, Dana M.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mastrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-915A-39

Alignment Scores:

Pred. No.: 1.52e-24
Score: 253.00
Percent Similarity: 87.10%
Best Local Similarity: 75.81%
Query Match: 74.19%
DB: 4

Length: 747
Matches: 47
Conservative: 7
Mismatches: 8
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-630-915A-39 (1-747)

QY 3 GlnlleaaglnValllleAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAla 22
Db 334 GAGATTGCTCAGTAACTTCACTGATATGCTTCTGCTGTTCTGAACAACCTTAGCCTTGCA 393

QY 23 ProGluGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 42
Db 394 CCAGGACAGTTAAATTAATTAATTCCTAAAGAAAATACAACTGGGTGGTGCAGAGAGAGTTA 453

QY 43 GlnlaaArgGlyLysLysArgGlnlleGlyTrrPheProAlaAsnTyrValIysLeuLeu 62
Db 454 CAGGCCAGAGAGAAAAAGCCGACAGAAAGAGATGTTTCTGCGACATGATTAACCTTTTG 513

QY 63 SerPro 64
Db 514 GGTCCA 519

RESULT 2

US-08-630-915A-193

Sequence 193, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630.915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:
LENGTH: 2873 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-630-915A-193

Alignment Scores:

Pred. No.: 8.64e-24
Score: 253.00
Percent Similarity: 87.10%
Best Local Similarity: 75.81%
Query Match: 74.19%
DB: 4

Length: 2873
Matches: 47
Conservative: 7
Mismatches: 8
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-630-915A-193 (1-2873)

QY 3 GlnlleaaglnValllleAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAla 22
Db 1119 GAGATTGCTCAGTAACTTCACTGATATGCTTCTGCTGTTCTGAACAACCTTAGCCTTGCA 1178

QY 23 ProGluGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 42
Db 1179 CCAGGACAGTTAAATTAATTAATTCCTAAAGAAAATACAACTGGGTGGTGCAGAGAGAGTTA 1238

QY 43 GlnlaaArgGlyLysLysArgGlnlleGlyTrrPheProAlaAsnTyrValIysLeuLeu 62
Db 1239 CAGGCCAGAGAGAAAAAGCCGACAGAAAGAGATGTTTCTGCGACATGATTAACCTTTTG 1298

QY 63 SerPro 64
Db 1299 GGTCCA 1304

RESULT 3

US-08-475-894-3

Sequence 3, Application US/08475894
Patent No. 5641748
GENERAL INFORMATION:

APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1975

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475.894
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both

Qy		GlucInleuIleAlaGlnValIleAlaSerTYrThrAlaTrnGlyProGluGlnLeuThreU	27
Dd		754 AAGATTACTCGCAAAAGTATATTTCCATATGTAGGCACAGAATGATGAATTCACAATC	8133
Qy		22 AlaProGlyGlnLeuIleLeuIleArgIlysAsn-----ProGlyGlyTrpTrpGlu	39
Dd		814 AAAGAAGGAGATVATCTACTCTCATCAATAAGACATGACATGACGTGGCTGGAGAA	8737
Qy		40 GIGlInleuGlnAlaArgIlyLysAsnGlnIleGlyTrpPheProAlaAsnTyVal	59
Dd		874 GGAGAGCTGAACGGACAGCA-----GGCGGTTCGCCGATTAACCTGTG	918
Qy		60 LysIleuLeuSerPro 64	
Dd		919 AAGTTACTTCACCG 933	
<hr/>			
RESULT 6			
	US-08-474-697-3	Sequence 3, Application US/08474697	
		Patent No. 6171800	
		GENERAL INFORMATION:	
		APPLICANT: Yen-Ming Hsu	
		TITLE OF INVENTION: THE CATP-LIKE GENE FAMILY	
		NUMBER OF SEQUENCES: 6	
		CORRESPONDENCE ADDRESS:	
		ADDRESSEE: LAHIVE & COCKFIELD	
		STREET: 60 State Street, Suite 510	
		CITY: Boston	
		STATE: Massachusetts	
		COUNTRY: USA	
		ZIP: 02109-1875	
		COMPUTER READABLE FORM:	
		MEDIUM TYPE: Floppy disk	
		COMPUTER: IBM PC compatible	
		OPERATING SYSTEM: PC-DOS/MS-DOS	
		SOFTWARE: PatentIn Release #1.0, Version #1.25	
		CURRENT APPLICATION DATA:	
		APPLICATION NUMBER: US/08/474,697	
		FILING DATE:	
		CLASSIFICATION: 435	
		ATTORNEY/AGENT INFORMATION:	
		NAME: Louis Myers	
		REGISTRATION NUMBER: 35,965	
		REFERENCE/DOCKET NUMBER: BGP-193	
		TELECOMMUNICATION INFORMATION:	
		TELEPHONE: (617)227-7400	
		TELEFAX: (617)227-5941	
		INFORMATION FOR SEQ ID NO: 3:	
		SEQUENCE CHARACTERISTICS:	
		LENGTH: 1392 base pairs	
		TYPE: nucleic acid	
		STRANDEDNESS: both	
		TOPOLOGY: linear	
		MOLECULE TYPE: cDNA	
		US-08-474-697-3	
<hr/>			
	Alignment Scores:		
	Pred. No.:	9.52e-06	Length: 1392
	Score:	113.50	Matches: 25
	Percent Similarity:	55.38%	Conservative: 11
	Best Local Similarity:	38.46%	Mismatches: 22
	Query Match:	33.28%	Indels: 7
		Gaps: 3	2
<hr/>			
	US-09-720-934-2_COPY_999_1062 (1-64) x US-08-474-697-3 (1-1392)		
Qy		2 GluGluIleAlaGlnValIleAlaSerTYrThrAlaTrnGlyProGluGlnLeuThreU	21
Dd		754 AAGATTACTCGCAAAAGTATATTTCCATATGTAGGCACAGAATGATGAATTCACAATC	8133
Qy		22 AlaProGlyGlnLeuIleLeuIleArgIlysAsn-----ProGlyGlyTrpTrpGlu	39
Dd		814 AAAGAAGGAGATVATCTACTCTCATCAATAAGACATGACATGACGTGGCTGGAGAA	8737

```

QY      40 GAGGAGCTGAACGGCAGCGCA-----GCCTGTTCCTCCCATTAATTGCTG   918
Db      874 GAGGAGCTGAACGGCAGCGCA-----GCCTGTTCCTCCCATTAATTGCTG   918
QY      60 LysleuleuserPro 64
Db      919 AAGTTACTTCCACC GG 933

RESULT 7
US-08-475-894-1
; Sequence 1, Application US/08475894
; Patent No. 5641748
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,894
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-475-894-1

Alignment Scores:
Pred. No.:          1,19e-05           Length:          1659
Score:              113.50             Matches:            25
Percent Similarity: 55.38%             Conservative:       11
Best Local Similarity: 38.46%           Mismatches:         22
Query Match:        33.28%             Indels:              7
DB:                  1                 Gaps:                2
US-09-720-934-2_COPY_999_1062 (1-64) x US-08-475-894-1 (1-1659)

QY      2 GluGluIleAlaGlnValIleAlaSerTyrThrAlaIhrGlyProGluGlnIleuThrLeu 21
Db      1021 AAGATTAATCTGCCAAGAATATTTTCATATGACGACAGAAATGATGATGATGACATC 1090
QY      22 AlaProGlyIleuIleuIleuIleuIleuArgLysLysAsn-----ProGlyLysTyrPTrGlu 39
Db      1081 AAAGAGAGAGATATAGTCACCTCATATCAATATAGACATGCATCGACGTATGCTGTGGGAA 1140
QY      40 GlyGluIeuGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrVal 59
Db      1141 GGAAGAGCTGAACGGCAGCGCA-----GGCGTGTTCCCCCGATTAACCTTGCTG 1185
QY      60 LysleuleuserPro 64

```

Db 1186 AAGTACTTCACCG 1200

RESULT 8

US-08-484-710-1

Sequence 1, Application US/08484710

Patent No. 5656438

GENERAL INFORMATION:

APPLICANT: Yen-Ming Hsu

TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,710

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Louis Myers

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: BGP-190

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1659 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-484-710-1

Alignment Scores:

Pred. No.:	1,19e-05	Length:	1659
Score:	113.50	Matches:	25
Percent Similarity:	55.38%	Conservative:	11
Best Local Similarity:	38.46%	Mismatches:	22
Query Match:	33.28%	Indels:	7
DB:	1	Gaps:	2

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-484-710-1 (1-1659)

QY 2 GluGluLeuAlaGlnValIleAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeu 21

Db 1021 AAGGATTACGCAAGATATATTTCCATATGAGGACAGAAATGATGATGATGACATC 1080

QY 22 AlaProGlyGlnLeuIleLeuIleArgLysLysAsn-----ProGlyGlyTyrPThrGln 39

Db 1081 AAGAAGAGATATAGCTCATCATCATATGAGACTGCATGCACTGAGCTGTGGGAA 1140

QY 40 GlyGluLeuGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaSerTyrVal 59

Db 1141 GGAGAGCTGAACGGACACACA-----GGCGTGTCCCGATTAATTGCTG 1185

QY 60 LysLeuLeuSerPro 64

Db 1186 AAGTACTTCACCG 1200

RESULT 9

US-08-484-709-1

Sequence 1, Application US/08484709

Patent No. 5837844

GENERAL INFORMATION:

APPLICANT: Yen-Ming Hsu

TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,709

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Louis Myers

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: BGP-192

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1659 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-484-709-1

Alignment Scores:

Pred. No.:	1,19e-05	Length:	1659
Score:	113.50	Matches:	25
Percent Similarity:	55.38%	Conservative:	11
Best Local Similarity:	38.46%	Mismatches:	22
Query Match:	33.28%	Indels:	7
DB:	2	Gaps:	2

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-484-709-1 (1-1659)

QY 2 GluGluLeuAlaGlnValIleAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeu 21

Db 1021 AAGGATTACGCAAGATATATTTCCATATGAGGACAGAAATGATGATGATGACATC 1080

QY 22 AlaProGlyGlnLeuIleLeuIleArgLysLysAsn-----ProGlyGlyTyrPThrGln 39

Db 1081 AAGAAGAGATATAGCTCATCATCATATGAGACTGCATGCACTGAGCTGTGGGAA 1140

QY 40 GlyGluLeuGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaSerTyrVal 59

Db 1141 GGAGAGCTGAACGGACACACA-----GGCGTGTCCCGATTAATTGCTG 1185

QY 60 LysLeuLeuSerPro 64

Db 1186 AAGTACTTCACCG 1200

RESULT 10

US-08-474-697-1

Sequence 1, Application US/08474697

Patent No. 6171800

GENERAL INFORMATION:

APPLICANT: Yen-Ming Hsu

TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, Suite 510

CITY: Boston

```
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,697
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGR-193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-474-697-1

Alignment Scores:
Pred. No.: 1,19e-05
Score: 113.50
Percent Similarity: 55.38%
Best Local Similarity: 38.46%
Query Match: 33.28%
DB: 3

Length: 1659
Matches: 25
Conservative: 11
Mismatch: 22
Indels: 7
Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-474-697-1 (1-1659)
QY 2 GluGluIleAlaGlnValIleAlaSerThrAlaThrGlyProGlnGlnLeuThrLeu 21
DB 1021 AAGGATTACTGCAAGATATATTTCCATATGAGCAGCAAGATGATGATGACATC 1080
QY 22 AlaProGlyGlnLeuIleLeuIleArgLysLysAsn-----ProGlyGlyTrpTrpGlu 39
DB 1081 AAAGAGAGATATATGATCTCTCATCAATAAGCAGCTGACATGAGCTGGGGGAA 1140
QY 40 GlyGluLeuGlnAlaArgGlyLysLysArgGlnIleGlyTrpPheProAlaAsnTrpVal 59
DB 1141 GGAGAGCTGAACGGCAGACGA-----GGCGTGTTCCCGATTAACCTTCGTG 1185
QY 60 LysLeuLeuSerPro 64
DB 1186 AAGTTACTTCACCG 1200

RESULT 11
US-08-671-354-1
; Sequence 1, Application US/08671354
; Patent No. 6423824
; GENERAL INFORMATION:
; APPLICANT: Hsu, Yen-Ming
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
```

```
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/671,354
FILING DATE: 27-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,344
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/484,709
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/475,894
FILING DATE: 07-JAN-1995
APPLICATION NUMBER: US 08/475,710
FILING DATE: 07-JAN-1995
APPLICATION NUMBER: US 08/474,697
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10274/009005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3143 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 436...2412
US-08-671-354-1

Alignment Scores:
Pred. No.: 2,73e-05
Score: 113.50
Percent Similarity: 55.38%
Best Local Similarity: 38.46%
Query Match: 33.28%
DB: 4

Length: 3143
Matches: 25
Conservative: 11
Mismatch: 22
Indels: 7
Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-671-354-1 (1-3143)
QY 2 GluGluIleAlaGlnValIleAlaSerThrAlaThrGlyProGlnGlnLeuThrLeu 21
DB 1222 AAGGATTACTGCAAGATATATTTCCATATGAGCAGCAAGATGATGATGACATC 1281
QY 22 AlaProGlyGlnLeuIleLeuIleArgLysLysAsn-----ProGlyGlyTrpTrpGlu 39
DB 1282 AAAGAGAGATATATGATCTCTCATCAATAAGCAGCTGACATGAGCTGGGGGAA 1341
QY 40 GlyGluLeuGlnAlaArgGlyLysLysArgGlnIleGlyTrpPheProAlaAsnTrpVal 59
DB 1342 GGAGAGCTGAACGGCAGACGA-----GGCGTGTTCCCGATTAACCTTCGTG 1386
QY 60 LysLeuLeuSerPro 64
DB 1387 AAGTTACTTCACCG 1401

RESULT 12
US-08-306-691B-48
; Sequence 48, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorina & Monaco, P.C.
```

```

1 STREET: Two Penn Center, Suite 1800
2 CITY: Philadelphia
3 STATE: Pennsylvania
4 COUNTRY: U.S.A.
5 ZIP: 19102
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
9
10 COMPUTER: IBM PS/2
11 OPERATING SYSTEM: MS-DOS
12 SOFTWARE: Wordperfect 5.1
13 CURRENT APPLICATION DATA:
14 APPLICATION NUMBER: US/08/306.691B
15 FILING DATE: September 15, 1994
16
17 CLASSIFICATION: 514
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER:
20 FILING DATE:
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Monaco, Daniel A.
23 REGISTRATION NUMBER: 30,480
24 REFERENCE/DOCKET NUMBER: 8321-8
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (215) 568-8383
27 TELEFAX: (215) 568-5549
28 TELE: No. 5734039e
29 INFORMATION FOR SEQ ID NO: 48:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 2757 base pairs
32 TYPE: nucleic acid
33 STRANDEDNESS: double
34 TOPOLOGY: linear
35
36 US-08-306-691B-48
37
38 Alignment Scores:
39 Pred. No.: 4.93e-05 Length: 2757
40 Score: 111.00 Matches: 21
41 Percent Similarity: 58.82% Conservative: 9
42 Best Local Similarity: 41.18% Mismatches: 15
43 Query Match: 32.55% Indels: 6
44 DB: 1 Gaps: 2
45
46 US-09-720-934-2_COPY_999_1062 (1-64) x US-08-306-691B-48 (1-2757)
47
48 QY 11 TyrThralathrglyProgluInleuThrDeuAlaProglyInleuile---leuile 29
49 ::::::::::::::::::::::::::::
50 Db 2343 TTCGCGCGCCGTGACCGTTGACAGCTGTGCTCAAGAGGGGTGACATCATCAAGATCCTT 2402
51 ::::::::::::::::::::::::::::::
52 QY 30 ArgylslysaenProglylyTyrTrrpGlueInlaArglyLyslysaArg 49
53 ::::::::::::::::::::::::::::::
54 Db 2403 AACAGAGAGGGACACGACGAGGCTGGGGGAGGAGATCATTCGCGG----- 2450
55 ::::::::::::::::::::::::::::::
56 QY 50 GlnileglyTrrpPheProAlaAsnTyrValys 60
57 ::::::::::::::::::::::::::::::
58 Db 2451 ---GTGGCTGTTCCTTCCTCAACTACGTGGAG 2480
59 ::::::::::::::::::::::::::::::
60
61 RESULT 13
62 PCT-US93-06251-79
63 Sequence 79, Application PC/TUS9306251
64 GENERAL INFORMATION:
65 APPLICANT: Wickstrom, Eric and Rife, Jason P.
66 TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
67 NUMBER OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
68 NUMBER OF SEQUENCES: 93
69 CORRESPONDENCE ADDRESS:
70 ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
71 STREET: 400 Garden City Plaza
72 CITY: Garden City
73 STATE: NY
74 COUNTRY: USA
75 ZIP: 11530
76
77 COMPUTER READABLE FORM:
78 MEDIUM TYPE: Floppy disk
79 COMPUTER: IBM PC compatible

```

[illegible]

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 921-5301
TELEFAX: (609) 921-4526
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 14..2545
US-07-646-537B-1

Alignment Scores:
Pred. No.:          9.21e-05      Length:      2793
Score:              109.00        Matches:     21
Percent Similarity: 58.82%       Conservative: 9
Best Local Similarity: 41.18%    Mismatches:  15
Query Match:        31.96%       Indels:      6
DB:                  1            Gaps:         2

US-09-720-934-2_COPY_999_1062 (1-64) x US-07-646-537B-1 (1-2793)
Cy      11 TyThrAlarhgIyProgluGlnLeuThrlLeuAalProgluGlnLeuile--Leuile 29
           ::|||::               :|||::               :|||::
Db      2387 TTCGTGGCCCGGACAGGTGCGAACTGTCCTTAAGAGGGTGATCATCATCAAGATCCTC 2446
Cy      30 ArglyselysanPrrogllyGlyTyTrpGluluglyglnLeuGlnlaArgglyLylslysarG 49
           2447 AATAAGAAGGAGCACGAAGGCTGTGGGGGTGGGAGATCAAGGCCGG----- 2494
Db      50 GlnlleGIyTpPhpeProAlaasnTyValys 60
           |||||::               :|||::               :|||::
Cy      2495 ---ATCGGCTGTTCCCTTCTTAATCATGTGGAG 2524
           |||||::               :|||::               :|||::

RESULT 15
US-09-023-655-1356
Sequence 1356, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-D001 US

```

```

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650) 855-0555
? TELEFAX: (650) 845-4166
? INFORMATION FOR SEQ ID NO: 1356:
? SEQUENCE CHARACTERISTICS:
?   LENGTH: 1414 base pairs
?   TYPE: nucleic acid
?   STRANDEDNESS: single
?   TOPOLOGY: linear
? IMMEDIATE SOURCE:
?   LIBRARY: GENBANK
?   CLONE: g35014
?   US-09-023-655-1356

Alignment Scores:
Pred. No.:      6.04e-05      Length:      1414
Score:          107.50       Matches:      20
Percent Similarity: 54.55%    Conservative: 10
Best Local Similarity: 36.36%  Mismatches:  20
Query Match:     31.52%      Indels:      5
DB:              4          Gaps:            1

US-09-720-934-2_COPY_999_1062 (1-64) x US-09-023-655-1356 (1-1414)
QY      5  AlAGlIvAlIlleAlaserYrThraIathrgIyProGInGInleuNrIleuAlaProGly 24
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      432  GCTTAATGTGAATTTTACTACATGCGCTGAGAGAGAGATGATTAATTCATTATGATTAAGGGG 491
QY      25  GInleuIlleuIlleuIleArGIsLysAsnProGlyGlyTPPTPGInGlyGInleuGlnAla 44
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      492  ACAAAAGTGATCGTCATGAGAAATGCAGTGCATGGTGGTGGCGTGGAGCTACATGGA 551
QY      45  ArgGlyLyLysLysArgGlnIleGlyITrPheProAlaSerIlyrval 59
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      552  -----CAAGTGGATGGATTCCCTTCAAACTATGTA 581

Search completed: July 1, 2004, 19:53:56
Job time : 49.815 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 196.347 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_999_1062
Perfect score: 341
Sequence: 1 GEEIAQVIASVTATGDEQLT.....RSKKRQIGFNPANYKLSP 64

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-Q/cgn2_1/USPTO.spool.p/US09720934/rnat 30062004 064541 13481/app query.fasta_1.1386
-DB=Published Applications NA -GEMT=fastap -SUFFIX=rmpb -MIMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USERS=US09720934 @CGN 1 1 1500 @rnat 30062004 064541 13481
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBIOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-RCAPOP=6 -RCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	319	93.5	2067	16	US-10-264-049-887	Sequence 887, App
2	253	74.2	747	9	US-09-879-957-39	Sequence 39, Appl
3	253	74.2	2873	9	US-09-879-957-193	Sequence 193, App
4	253	74.2	3746	11	US-09-764-875-176	Sequence 176, App
5	253	74.2	4053	13	US-10-342-887-1882	Sequence 1882, Ap
6	253	74.2	4053	13	US-10-172-118-1882	Sequence 1882, Ap
7	253	74.2	4210	9	US-09-764-868-125	Sequence 125, App
8	253	74.2	5828	13	US-10-398-885A-15	Sequence 15, Appl
9	117.5	34.5	852	15	US-10-259-165-471	Sequence 471, App
10	117.5	34.5	855	15	US-10-259-165-143	Sequence 143, App
11	117.5	34.5	1563	17	US-10-437-963-41013	Sequence 41013, A
12	116.5	34.2	564	15	US-10-029-386-15	Sequence 15, Appl
13	114.5	33.6	361	13	US-10-425-114-5871	Sequence 5871, Ap
14	114.5	33.6	1420	13	US-10-425-114-5871	Sequence 31465, A
15	114.5	33.6	1514	13	US-10-425-114-25238	Sequence 25238, A
16	114.5	33.6	1514	13	US-10-425-114-28939	Sequence 28939, A
17	114.5	33.6	1559	13	US-10-085-783A-37196	Sequence 25722, A
18	113.5	33.3	386	13	US-10-085-783A-37196	Sequence 37196, A
19	113.5	33.3	386	16	US-10-242-535A-37196	Sequence 37196, A
20	113.5	33.3	400	13	US-10-085-783A-22639	Sequence 22639, A
21	113.5	33.3	400	16	US-10-242-535A-22639	Sequence 22639, A
22	113.5	33.3	1215	17	US-10-648-593-61	Sequence 61, Appl
23	113.5	33.3	3143	14	US-10-144-621-1	Sequence 1, Appl
24	113.5	33.3	3348	13	US-10-342-887-187	Sequence 187, App
25	113.5	33.3	3348	13	US-10-172-118-187	Sequence 187, App
26	112.5	33.0	1582	15	US-10-259-165-580	Sequence 580, App
27	112.5	33.0	1661	13	US-10-425-114-34107	Sequence 34107, A
28	111	32.6	518	9	US-09-796-692-5870	Sequence 5870, Ap
29	111	32.6	518	15	US-10-040-862-5870	Sequence 5870, Ap
30	111	32.6	518	16	US-10-057-475B-5870	Sequence 5870, Ap
31	111	32.6	518	16	US-10-154-884B-5870	Sequence 5870, Ap
32	111	32.6	2757	10	US-09-960-706-1059	Sequence 1059, Ap
33	109.5	32.1	601	15	US-10-259-165-494	Sequence 494, App
34	107.5	31.5	1380	13	US-10-087-192-140	Sequence 140, App
35	107.5	31.5	1414	13	US-10-342-887-1250	Sequence 1250, Ap
36	107.5	31.5	1414	13	US-10-172-118-1250	Sequence 1250, Ap
37	107.5	31.5	1414	17	US-10-641-643-1356	Sequence 1356, Ap
38	107.5	31.5	1513	15	US-10-210-120-20	Sequence 20, Appl
39	107.5	31.5	3227	16	US-10-115-831-10	Sequence 10, Appl
40	107.5	31.5	79831	13	US-10-087-192-139	Sequence 139, App
41	106.5	31.2	909	13	US-10-425-114-9122	Sequence 9122, Ap
42	106.5	31.2	971	9	US-09-879-957-197	Sequence 197, App
43	106.5	31.2	1334	13	US-10-087-192-143	Sequence 143, App
44	106.5	31.2	1603	13	US-10-424-599-11257	Sequence 131257, Sequence 123461,
45	106.5	31.2	1890	13	US-10-424-599-123461	

ALIGNMENTS

RESULT 1
US-10-264-049-887
; Sequence 887, Application US/10264049
; Publication No. US2004000579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL33PI
; CURRENT APPLICATION NUMBER: US/10/264, 049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 887
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)

OTHER INFORMATION: n equals a,t,b, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (2058) ..(2058)
OTHER INFORMATION: n equals a,t,b, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (2063) ..(2063)
OTHER INFORMATION: n equals a,t,b, or c
US-10-264-049-887

Alignment Scores:
Pred. No.: 3,35e-35
Score: 319.00
Percent Similarity: 98.39%
Best Local Similarity: 95.16%
Query Match: 93.55%
Length: 2067
Matches: 59
Conservative: 2
Mismatch: 1
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-264-049-887 (1-2067)

QY 3 Glu1lea1aGlnVal1leAlaSerTyrThrAlaThrGlyProGluGlnLeuThrlaAla 22
Db 62 AAATGGCCCAAGGTATTCCTCATACACCGCCCGCCGAGCAGCTCCTCGCC 121

QY 23 ProGlyGlnLeu1leAlaSerTyrThrAlaThrGlyProGluGlnLeuThrlaAla 42
Db 122 CCGTGTACGCTGATTTGATTCGAAAAAGAACCCAGGTGATGTGGAGAGAGCTG 181

QY 43 Gln1laa1aGlyLysLysArgGln1leGlyTrrPheProAlaSerTyrValLysLeuLeu 62
Db 182 CAAGCAGGTGGAAAAAGCGCCAGATAGGCTGTCCAGCTATTTATGTAACTTCTA 241

QY 63 SerPro 64
Db 242 AGCCT 247

RESULT 2
US-09-879-957-39
Sequence 39, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
HOFMAN, No. US20020034755A1h
KAY, Brian K.
FOWLES, Dana M.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/630,915
FILING DATE: 03-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-879-957-39

Alignment Scores:
Pred. No.: 2,32e-26
Score: 253.00
Percent Similarity: 87.10%
Best Local Similarity: 75.81%
Query Match: 74.19%
Length: 747
Matches: 47
Conservative: 7
Mismatch: 8
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x US-09-879-957-39 (1-747)

QY 3 Glu1lea1aGlnVal1leAlaSerTyrThrAlaThrGlyProGluGlnLeuThrlaAla 22
Db 334 GAGATTCCTCAGTAACCTTCATATGCTTCTGTAACAACCTTAGCCTTGA 393

QY 23 ProGlyGlnLeu1leAlaSerTyrThrAlaThrGlyProGluGlnLeuThrlaAla 42
Db 394 CCAAGCAGTTAATTAATTCTAAGAAAAATCAAGTGGTGTGTGGCAAGGAGTTA 453

QY 43 Gln1laa1aGlyLysLysArgGln1leGlyTrrPheProAlaSerTyrValLysLeuLeu 62
Db 454 CAGGCCAGAGAAAAAGCAGCAAGAAAGATGTTCTCCAGCTCATGTTAACTTTG 513

QY 63 SerPro 64
Db 514 GGTCCA 519

RESULT 3
US-09-879-957-193
Sequence 193, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
HOFMAN, No. US20020034755A1h
KAY, Brian K.
FOWLES, Dana M.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/630,915


```

Db      3124  AAGATGTCGACGTAACCTTCAGCAATATGTTCTTCTGATTCTGAACAACCTAGCCTTGCA 31183
Oy      23    ProglglnLeuIleuIleArglylsAsnProglgylYTrpTpgLuglyLeu 42
Db      3184  CCAGGACAGTTAAATATTAATCTAAAGAAAATCTCAAGTGGGTGGTGGCAAGGAGATTGA 3243
Oy      43    GlnAlaArgglylsYlsYlsArgGlnIleGlyTrpPheProIlaenYrValylsLeu 62
Db      3244  CAGGCGCAAGAGAAAAAAGCAGCAAGAAAGATGGTTTCTGCGACGTCAATGTTAAACTTTTG 3303
Oy      63    SerPro 64
Db      3304  GGTCCA 3309

RESULT 9
US-10-259-165-471
; Sequence 471, Application US/10259165
; Publication No. US20030135888A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hur-song
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-1NP
; CURRENT APPLICATION NUMBER: US/10/259,165
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/368,327
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: Patentlast.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 471
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (391)..(391)
; OTHER INFORMATION: n = any nucleotide
US-10-259-165-471

Alignment Scores:
Pred. No.: 4.47e-07 Length: 852
Score: 117.50 Matches: 25
Percent Similarity: 59.65% Conservative: 9
Best Local Similarity: 43.86% Mismatches: 18
Query Match: 34.46% Indels: 5
DB: 15 Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-259-165-471 (1-852)
Oy      4    IleAlaGlnValIleIleAlaSerYrThrAlaThrGlyProgluGlnLeuThrLeuAlaPro 23
Db      655  TTGGCTGAGGCCATCCAGTCTTATGCTGCTAGAGAGTGAACCTGAAGCTCAACTGGCAGCT 714
Oy      24    GlyGlnLeuIleLeuIleArglylsAsnProglgylYTrpTpgLuglyLeuGln 43
Db      715  GGTGACTATATATGTTGTCGGAAGGTGTCAAAACAATGATGGCGAAGGTGA----- 768
Oy      44    AlaArgglylsYlsYlsArgGlnIleGlyTrpPheProIlaenYrValyls 60

```

```

Db      769  TGCAGAGGAGAA-----GCTGGCTGGTTCCTTACGACTACATCGAG 810

RESULT 10
US-10-259-165-143
; Sequence 143, Application US/10259165
; Publication No. US20030135888A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hun-song
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-NP
; CURRENT APPLICATION NUMBER: US/10/259,165
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/368,327
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: Patentlist.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 143
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (391)..(391)
; OTHER INFORMATION: n = any nucleotide
US-10-259-165-143

Alignment Scores:
Pred. No.:      4.49e-07      Length:      855
Score:          117.50      Matches:      25
Percent Similarity: 59.65%      Conservative: 9
Best local Similarity: 43.86%      Mismatches:  18
Query Match:    34.46%      Indels:      5
DB:             15          Gaps:          2

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-259-165-143 (1-855)
Qy      4  I::l::a::n::v::l::l::e::a::s::e::r::T::r::T::h::A::l::a::t::h::r::G::l::y::P::r::o::g::l::u::n::l::e::u::l::a::p::r::o 23
Db      655  TTGGCTGAGGCGCATCCAGCTCTTATCGCGCTCGAGAGTGAACATGAGCTCAACCTCGCGAGCT 714
Qy      24  G::l::y::l::e::u::l::l::e::u::l::e::a::r::g::l::y::l::a::s::n::p::r::o::g::l::y::l::T::r::T::p::g::l::u::l::e::u::l::n 43
Db      715  GGTGACTATATAGTTGTCGGAAGGTGTCAACATGAGTGGCGAGAAAGGTGAA----- 768
Qy      44  A::l::a::r::g::l::y::l::y::s::a::r::g::l::l::l::e::g::l::T::r::P::h::e::r::o::l::a::s::e::n::T::r::V::a::l::y::s 60
Db      769  TGCAGAGGAGAA-----GCTGGCTGGTTCCTTACGACTACATCGAG 810

RESULT 11
US-10-437-963-41013
; Sequence 41013, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

```

```

APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 41013
LENGTH: 1563
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_4439C.1
US-10-437-963-41013

Alignment Scores:
Pred. No.: 9,12e-07
Score: 117.50
Percent Similarity: 59.65%
Best Local Similarity: 43.86%
Query Match: 34.46%
DB: 17

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-437-963-41013 (1-1563)

QY 4 11ealaglinallealaserYrThralaThrglyProgluglnleuthrleuAlaPro 23
DB 1093 TTGGCTAGGCGACATCCAGCTCTTACGTCGAGAGTGAAACTGCAACCTGCGAGCT 1152

QY 24 G1yGlnleulleuileuileaglylsAsnProglylYrTpPgluglYglnleuGln 43
DB 1153 GGTCACATATAGTGTGTCGGAGAGCTGTCAACATGATGGCGAGAGTGAA----- 1206

QY 44 AlaArgGlylsYlsAArgGlnlleglYrTpPheProAlaAsnYrValys 60
DB 1207 TGCAGAGGAGAA-----GCTGGCTGGTTCCTTACGACTACATCGAG 1248

RESULT 12
US-10-029-386-15/C
Sequence 15, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AEWOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 15
LENGTH: 564
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL135783.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.55
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.52
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.51
OTHER INFORMATION: SWISSPROT HIT Q14155, EVALUE 6.00e-27
OTHER INFORMATION: NT HIT: g14745025, EVALUE 1.00e-110
OTHER INFORMATION: EST_HUMAN HIT F20693.1, EVALUE 0.00e+00
US-10-029-386-15

Alignment Scores:
Pred. No.: 3.82e-07
Score: 116.50
Percent Similarity: 58.82%
Length: 564
Matches: 19
Conservative: 11

```

```

Best Local Similarity: 37.25%
Query Match: 34.16%
DB: 15
Gaps: 1

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-029-386-15 (1-564)

QY 10 serYrThrAlaThrglyProgluglnleuthrleuAlaProglYglnleulle 29
DB 408 AACTTAAAGCAGACTAATGAGATGAAGTGTGAGGAGGACATTAATGAGAG 349

QY 30 ArgYlsYlsAsnProglYglnYrTpPgluglYglnleuGlnAlaArgGlylsA 49
DB 348 ACAGGAGTTGAAGAGAGGCTGTGGGAGGCACTTAATATGGAGAG----- 301

QY 50 GlnlleglYrTpPheProAlaAsnYrValys 60
DB 300 ---ACAGGCTGGTTCCTTACGACTAATATGTCCT 271

RESULT 13
US-10-425-114-5871
Sequence 5871, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53113)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 5871
LENGTH: 361
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700476119_FLI
US-10-425-114-5871

Alignment Scores:
Pred. No.: 4.35e-07
Score: 114.50
Percent Similarity: 59.65%
Best Local Similarity: 42.11%
Query Match: 33.58%
DB: 13
Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-425-114-5871 (1-361)

QY 4 11ealaglinallealaserYrThralaThrglyProgluglnleuthrleuAlaPro 23
DB 36 TTGGCTAGGCGAATCCAGCTCATACCAAGCCGAAGCATACCGAGCTCAACCTTCAACT 95

QY 24 G1yGlnleulleuileuileaglylsAsnProglylYrTpPgluglYglnleuGln 43
DB 96 GGTCACATATAGTGTGTCGGAGAGCTGTCAACATGATGGCGGAGAGTGAA----- 149

QY 44 AlaArgGlylsYlsAArgGlnlleglYrTpPheProAlaAsnYrValys 60
DB 150 TGCAGAGGAGAA-----GCTGGTGGTTCCTTACGACTACATCGAG 191

RESULT 14
US-10-425-114-31465
Sequence 31465, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.

```

```
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ. ID NOS: 73128
/ SEQ. ID NO 31465
/ LENGTH: 1420
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZMFLB73177F01_FLI
US-10-425-114-31465

Alignment Scores:
Pred. No.: 2,17e-06 Length: 1420
Score: 114.50 Matches: 22
Percent Similarity: 57.89% Conservative: 11
Best Local Similarity: 38.60% Mismatches: 19
Query Match: 33.58% Indels: 5
DB: 13 Gaps: 1

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-425-114-31465 (1-1420)
QY 4 11eAlaGlnVal11leAlaSerTyrThrAlaThrGlyProGlnLeuThrLeuAlaPro 23
DB 924 CTGGCTGAGGCAATCCATCATATCGAGCAGACGACACGTGAGCTCAACCTTTCAGCT 983
QY 24 G1yGlnLeu11leu1leaArgLysLysAsnProG1yG1yTrrPrrGlnG1uLeuGln 43
DB 984 GGTGACTACATATAGTGTCCGAAAGGTGTGCAACAAATGAGTGGCCGAAAGGTGAATGCAGG 1043
QY 44 AlaArgG1yLysLysArgGln11leG1yTrrPrrPheProAlaAsnTyrValLys 60
DB 1044 GGGAGAGCT-----GGCTGTGTTCCCTACGACTACATCGAG 1079

RESULT 15
US-10-425-114-25238
/ Sequence 25238, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ. ID NOS: 73128
/ SEQ. ID NO 25238
/ LENGTH: 1505
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB732-023-C2_FLI
US-10-425-114-25238

Alignment Scores:
Pred. No.: 2.32e-06 Length: 1505
Score: 114.50 Matches: 24
Percent Similarity: 59.65% Conservative: 10
Best Local Similarity: 42.11% Mismatches: 18
Query Match: 33.58% Indels: 5
DB: 13 Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-425-114-25238 (1-1505)
```

```
QY 4 11eAlaGlnVal11leAlaSerTyrThrAlaThrGlyProGlnLeuThrLeuAlaPro 23
DB 1115 TTGGCTGAGGCAATCCATCATATCCGAAAGCGAATACCGAGCTCAACCTTTCAGCT 1174
QY 24 G1yGlnLeu11leu1leaArgLysLysAsnProG1yG1yTrrPrrGlnG1uLeuGln 43
DB 1175 GGTGACTACATATAGTGTCCGAAAGGTGTGCAACAAATGAGTGGCCGAAAGGTGAATGCAGG 1228
QY 44 AlaArgG1yLysLysArgGln11leG1yTrrPrrPheProAlaAsnTyrValLys 60
DB 1229 TGCAGGGGGGAAA-----GCTGTGTTGTTCCCTACGACTACATCGAG 1270
```

Search completed: July 1, 2004, 20:16:08
Job time : 202.347 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 1925.25 Seconds

(without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_999_1062
Perfect score: 341
Sequence: 1 GEEIQVASTATATGEEQT.....RQKKQIGMFANYKLLSP 64

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgm2_1/USPTO.spool.p/US09720934/runat 30062004 064540 13442/app query.fasta_1.1386
-DB=EST -QFMT=fastap -SUFFIX=est -MISMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720934.@CGN 1 1.12421 @runat 30062004 064540 13442 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST: *
1: em_escba: *
2: em_escbm: *
3: em_escin: *
4: em_escmu: *
5: em_escov: *
6: em_escpl: *
7: em_escro: *
8: em_escro: *
9: gb_esc1: *
10: gb_esc2: *
11: gb_esc3: *
12: gb_esc4: *
13: gb_esc5: *
14: gb_escfun: *
15: em_escfun: *
16: em_escro: *
17: em_escro: *
18: em_escro: *
19: em_escro: *
20: em_escro: *
21: em_escro: *
22: em_escro: *
23: em_escro: *
24: em_escro: *
25: em_escro: *
26: em_escro: *
27: em_escro: *
28: gb_esc1: *

29: gb_gses2: *
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	330	96.8	538	14	CA390605	CA390605 cs11d07
2	330	96.8	545	14	CB161912	CB161912 K-EST0221
3	330	96.8	583	14	CB156803	CB156803 K-EST0215
4	330	96.8	600	9	AL134506	AL134506 DKF2P547M
5	330	96.8	628	13	BQ393010	BQ393010 NISC mq28
6	330	96.8	629	14	CF768246	CF768246 CBS003921
7	330	96.8	750	14	CD356662	CD356662 AGENCOURT
8	330	96.8	1062	13	BQ898718	BQ898718 AGENCOURT
9	335	95.3	446	29	CG508946	CG508946 OST19780
10	325	95.3	537	9	AA063751	AA063751 mj79g10.y
11	325	95.3	603	9	AI594919	AI594919 mj79g10.y
12	325	95.3	610	13	BQ443095	BQ443095 UI-M-EVO-
13	325	95.3	694	13	BU054334	BU054334 UI-M-EVO-
14	325	95.3	732	12	BM950476	BM950476 UI-M-EVO-
15	325	95.3	782	13	BU614534	BU614534 UI-M-EVO-
16	325	95.3	875	10	BF122946	BF122946 601761612
17	325	95.3	898	14	CA987804	CA987804 AGENCOURT
18	325	95.3	906	13	BU519029	BU519029 AGENCOURT
19	325	95.3	2589	11	AK082606	AK082606 Mus muscu
20	325	95.3	5385	11	BC062938	BC062938 Mus muscu
21	318	93.3	575	14	CB163763	CB163763 K-EST0224
22	318	93.3	636	14	CD347867	CD347867 UI-M-FY0-
23	317	93.0	633	14	CA353372	CA353372 624874 NC
24	317	93.0	760	13	BX879964	BX879964 BX879964
25	317	93.0	805	13	BU451981	BU451981 603767625
26	311	91.2	668	14	CF763755	CF763755 CBS005445
27	299	87.7	346	14	R39287	R39287 yte69g01.t1
28	286	83.9	522	28	AQ252614	AQ252614 HS_2010.B
29	268	78.6	298	14	F06100	F06100 HSC0XA031 n
30	265	77.7	518	28	AO7229475	AO7229475 HS_2154.B
31	256	75.1	590	9	AV886072	AV886072 AV886072
32	256	75.1	637	13	BM232778	BM232778 BM232778
33	256	75.1	709	9	AV849639	AV849639 AV849639
34	256	75.1	712	13	BM092285	BM092285 BM092285
35	253	74.2	362	12	BG089231	BG089231 uc64d07.y
36	253	74.2	463	13	BX504944	BX504944 DKF2P686G
37	253	74.2	1503	11	AK054527	AK054527 Mus muscu
38	252	73.9	785	13	BX276672	BX276672 BX276672
39	249	73.0	652	14	CF766841	CF766841 CBS003922
40	248	72.7	274	14	F07724	F07724 HSC2FG081 n
41	244	71.6	443	13	BU430865	BU430865 UI-HF-BNO
42	243	71.3	764	13	BU106683	BU106683 60254522
43	237	69.5	418	14	CF144063	CF144063 UI-HF-BPO
44	236	69.2	309	9	AJ281124	AJ281124 4A3A-AB-
45	236	69.2	408	9	AJ283863	AJ283863 4A3B-AAG-

ALIGNMENTS

RESULT 1
LOCUS CA390605 538 bp mRNA linear EST 06-NOV-2002
DEFINITION cs11d07.y1 Human Retinal pigment epithelium/choroid cDNA
(Un-normalized, unamplified): cs Homo sapiens cDNA clone cs11d07
5', mRNA sequence.
ACCESSION CA390605
VERSION CA390605.1 GI:24721774
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1 (bases 1 to 518)	
AUTHORS	Wistow, G., Bernstein, S.L., Wyatt, M.K., Farris, R.N., Behal, A., Touchman, J.W., Bouffard, B., Smith, D., and Peterson, K.	
TITLE	Expressed sequence tag analysis of human RPE/choroid for the NEIBank Project: Over 600 non-redundant transcripts, novel genes and splice variants	
JOURNAL	Mol. Vis. 8 (4), 205-220 (2002)	
MEDLINE	22103460	
PUBMED	12107410	
COMMENT	Contact: Wistow G Section on Molecular Structure and Function National Eye Institute 6/331, NIH, Bethesda, MD 20892-2740, USA Tel: 301 402 3452 Fax: 301 496 0078 Email: graeme@helix.nih.gov Plate: 111 row d column: 07 Seq primer: M13RPI reverse primer (ABI). Location/Qualifiers	
FEATURES	<p>1..518</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="cs111607"</p> <p>/tissue_type="RPE/choroid"</p> <p>/dev_stage="Adult"</p> <p>/lab_host="EMDHL0B"</p> <p>/clone_lib="Human Retinal pigment epithelium/choroid cDNA (un-normalized, unamplified): cs"</p> <p>/note="Organ: Eye; Vector: pCMVSPORT; Two different donor eyes (75-80 years old) yielded approximately 600 mg of dissected RPE/choroid tissue. This in turn yielded 340 ug of total RNA and 7 ug of mRNA. A directionally cloned cDNA library in the pCMVSPORT vector was constructed at Life Technologies (Rockville, MD; now part of Invitrogen Corp), essentially following the protocols of the SuperScript Plasmid System (Invitrogen Corp</p> <p><http://www.invitrogen.com/>). The library code designation was cs. For this library, cDNA inserts were cloned into the NotI/MluI sites of the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."</p>	
ORIGIN		
Alignment Scores:		
Pred. No.:	2.27e-28	Length: 538
Score:	330.00	Matches: 62
Percent Similarity:	100.00%	Conservative: 0
Best local Similarity:	100.00%	Mismatches: 0
Query Match:	96.77%	Indels: 0
DB:	14	Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64)	x CA390605 (1-538)	
QY	3 GlltlelaaglnVallllealeasertYThralaThrglyProglnglnleuThrlenuA 22	
Db	131 GAATTCGCCAGGTATTATTCCTCATACACCCGCCACCGCCCCGAGCAGCTACTCTGGCC 190	
QY	23 ProglnglnleulleleulleaKGLyslysaSnProglnglyYrPTTProglnglyGlnleu 42	
Db	191 CTTGTAGAGCTGATTTTGATCCGAAAAAGAACCCAGGTGATGTTGGGAAGAGAGCTG 250	
QY	43 GlnlaaKProglnglylsYsaRglnlllglyYrPhpPoalaasnyYVallysleu 62	
Db	251 CAACCAAGCTGGGAAAAAGCGCCAAATAGAGCTGTTCACGCTAATTATGTAAAGCTTCTA 310	
QY	63 SerPro 64	
Db	311 AGCCCT 316	
RESULT 2		
CB161912	545 bp	mRNA linear EST 30-JAN-2003
LOCUS		

DEFINITION	K-EST0221969 L17N670205n1 Homo sapiens cDNA clone		
ACCESSION	L17N670205n1-41-H01 5', mRNA sequence.		
VERSION	CB161912		
KEYWORDS	EST.		
SOURCE	CB161912.1 GI:28148038		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,U.M., Park,H.S., Kim,S. and Kim,Y.S.		
TITLE	21C Frontier Korean EST Project 2001		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongjung@mail.krrib.re.kr Plate: 41 row: H column: 01 High quality sequence stop: 545. Location/Qualifiers 1..545 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="L17N670205n1-41-H01" /sex="F" /lab_host="Top10F" /clone_lib="L17N670205n1" /note="Organ: Liver; Vector: pT7TS-Pac; Site 1: Scori; Site 2: Notci; The library was constructed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.3e-28	length:	545
Score:	330.00	Matches:	62
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	96.77%	Indels:	0
DB:	14	Gaps:	0
US-09-720-934-2_COPY_999_1062 (1-64) x CB161912 (1-545)			
OY	3	GluUleAaGlnValIleAlaIleAsrTYrThAlaThrGlyProGluGlnLeuThrIleuAla	22
Db	124	GAATTCGCCCAAGGTATTGGCTCATACACCGCACCGGCCCGAGCAGCTCAGCTTCGCC	183
OY	23	ProGlyGlnLeuIleLeuIleLeuIleArgLyLyAsnProGlyGlyTrpTrpGluGlyGluLeu	42
Db	184	CTGGTACGCTGATTGGATCCGAAAAAAGAACCCAGGTGGATGGGAGAGAGAGCTG	243
OY	43	GlnAlaArgGlyLyLyLyArgGlnIleGlyTrpPheProAlaAsnYrValLySleuLeu	62
Db	244	CAACGACGTGGGAAAAAGCCGCGCAGATAGGCTGGTCCACGTAATATGTAAAGCTTCTA	303
OY	63	SerPro 64	
Db	304	AGCCTT 309	
RESULT 3			
LOCUS	CB156803 583 bp mRNA linear EST 29-JAN-2003		
DEFINITION	K-EST0215771 L17N670205n1 Homo sapiens cDNA clone		
ACCESSION	L17N670205n1-4-C07 5', mRNA sequence.		
VERSION	CB156803		
KEYWORDS	EST.		
SOURCE	CB156803		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,U.M., Park,H.S., Kim,S. and Kim,Y.S.		
TITLE	21C Frontier Korean EST Project 2001		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongjung@mail.krrib.re.kr Plate: 41 row: H column: 01 High quality sequence stop: 545. Location/Qualifiers 1..545 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="L17N670205n1-41-H01" /sex="F" /lab_host="Top10F" /clone_lib="L17N670205n1" /note="Organ: Liver; Vector: pT7TS-Pac; Site 1: Scori; Site 2: Notci; The library was constructed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."		

VERSION	CB156803.1	GI:28141931
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Bukacinsky, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 583)	
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.U., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.	
TITLE	21C Frontier Korean EST Project 2001	
JOURNAL	Unpublished (2002)	
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsung@mail.kribb.re.kr Plate: 4 row: C column: 07 High quality sequence stop: 583.	
FEATURES	Location/Qualifiers	
source	1..583 /organism="Homo sapiens" /mol_type="RNA" /db_xref="taxon:9606" /clone="L17N670205n1-4-C07" /sex="F" /lab_host="Trop10F" /clone_idb="L17N670205n1" /note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI, Site 2: NotI; The library was contributed by the Soares Laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6 (9): 791-806. RNA was prepared from harvested cell culture."	
ORIGIN		
Alignment Scores:		
Pred. No.:	2.5e-28	Length: 583
Score:	330.00	Matches: 62
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	96.77%	Indels: 0
DB:	14	Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x CB156803 (1-583)		
OY	3	Glul1ealaglnVal1le1a1seryThr1a1nrhgiyProgiuglnleuthre1ua1a 22
Db	124	GAATATGCCCGAGTATATGCTCATACACCGCCACCGCCCGAGCAGCTCAGCTCTGCC 183
OY	23	Progiyginleulle1eulle1e1aArgyys1ysa5npProgiygiYTPPTTg1uG1yG1u1eu 42
Db	184	CCTGGTCAGCTGATTTTATCCGAAAAAGAACCCAGGTGATGCTG3GAAGAGAGAGCTG 243
OY	43	G1na1aArg31yLs1yArG1ni1leG1yTPPhPePoa1aAs1ryVal1ys1euden 62
Db	244	CAAGAGCTGGGAAAAAGCCGACAGTACGCTGTTCCAGCTAA1TATGTAAGCTTCTA 303
OY	63	SerPeP 64
Db	304	AGCCCT 309
RESULT 4		
LOCUS	AL134506	600 bp mRNA linear EST 04-SEP-2003
DEFINITION	DKFZP547M065_r1 547 (synonym: hfpri1) Homo sapiens cDNA clone	
ACCESSION	DKFZP547M065.5	mRNA sequence.
KEYWORDS	AL134506	
VERSION	AL134506.1	GI:6602693
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 600)
AUTHORS Blum,H., Bauersachs,S., Mewes,W., Weil,B. and Wiemann,S.
TITLE EST (Blum,H., Bauersachs,S., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS

MIPS
IGPstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the CDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp547M065) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcententrum, Neuherweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp547M065"
/tissue_type="brain"
/dev_stage="fetal"
/lab_host="X1-2blue"
/clone_id="547 (synonym: hfbf1)"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

ALIGNMENT SCORES:
Pred. No.: 2.58e-28 Length: 600
Score: 330.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.77% Indels: 0
DB: 9 gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AL134506 (1-600)

QY 3 GluuileaglnValllleaIseryrThAlathrglyProgluglnLeutnrleuaA 22
Dd 260 GAATTGCCCAAGTATTCCTCATACACCGCCACGCGCCCCGACGACTCCTCGCC 319
QY 23 ProGlYglnleullelleulearGLysLYAsnProGlYslYTPPTPglnGLnuen 42
Dd 320 CCTGGTCAGCTGAATTTTGATCCGAAAAAAACAACCAGGTGGATGTGGGAAGAGAGCTG 379
QY 43 GlAlAlArGLylYslYSArGlnInleGLYTTrPherobAlasentYrVAlLYlsleuEn 62
Dd 380 CAAGCACGTGGAAAAGCGCCAGATAGGCTGTGTTCCAGCTAATTATGTAAGCTTCTA 439
QY 63 SerPro 64
Dd 440 AGCCCT 445

RESULT 5
LOCUS BQ393010
DEFINITION NISC mg28c06.y1 NICHD XGC Emb5 Silurana tropicalis cDNA clone
IMAGE:5384843 5', mRNA sequence.
ACCESSION BQ393010
VERSION BQ393010.1 GI:21080687
SOURCE EST.
ORGANISM Silurana tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Silurana tropicalis
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae, Piplidae;
Xenopodinae; Silurana.
1 (bases 1 to 628)

AUTHORS NIH-XGC <http://image.jnl.gov/image/html/xenopuslib.info.shtml>.
TITLE National Institute of Child Health and Human Development, National Cancer Institute, Xenopus Gene Collection
JOURNAL Unpublished (2002)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 cDNA Library Preparation: cDNA Library Arrayed by: The I.M.A.G.E. Consortium/ILNL
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
 Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.jnl.gov>
 Plate: ILMA1980 row: F column: 12
 Seq primer: M13RPI reverse primer (ABI).
FEATURES Location/Qualifiers
 source
 1..628
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:584843"
 /tissue_type="gastrula"
 /dev_stage="embryo, stages 10-13"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NICHD XGC Emb5"
 /note="Vector: PCMV-SPORE 1; Site 1: NotI; Site 2: EcoRV; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.0 Kb. Constructed by Invitrogen. Note: This is a Xenopus Gene Collection (XGC) library."
ORIGIN
 Alignment Scores:
 Pred. No.: 2.73e-28 Length: 628
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x BQ393010 (1-628)
 QY 3 GtllleaglnvalilleaaserythrAlaThrglyProgluglnleuThrlenuA 22
 |||||
 Db 416 GAATTCGCCAGGTATTGCTTTACACAGCAACTGGCCTGAGCAGCTGACACTGGCT 475
 QY 23 Proglglnleuileuileuileuileuileuileuileuileuileuileuileu 42
 |||||
 Db 476 CTTGGCAACTTATCTCATCTCGAAAAAGATCTCGAGAGATGCTGGAGAGAACTG 535
 QY 43 GlnAlaArgGlyLysLysArgGlnlleGlyTrpPheProAlaAsnTyrrValyleu 62
 |||||
 Db 536 CAGGCAAGTGGAGAAAGCAGCAAGATTGGATGTTTCCAGCAATATGTAAGCTTTG 595
 QY 63 SerPro 64
 |||||
 Db 596 AGCCCT 601
RESULT 6
LOCUS CF768246
DEFINITION C6S003921 Bos taurus skin cDNA library Bos taurus cDNA clone
 C6S003921 5', mRNA sequence.
ACCESSION CF768246
VERSION CF768246.1 GI:37717465
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.
REFERENCE 1 (bases 1 to 629)
AUTHORS Wang, Y.H., McWilliam, S., and Lehnert, S.
TITLE Transcription profiling of cattle skin

JOURNAL Unpublished (2003)
COMMENT Contact: Dr Yonghong Wang
 Functional Genomics Lab
 CSIRO Livestock Industries
 Level 5, Queensland Biosciences Precinct,
 306 Carmody Road St. Lucia QLD Australia
 Tel: 07 3214 2445
 Fax: 07 3214 2685
 Email: Yonghong.Wang@csiro.au
 Plate: 5 row: A column: 11.
FEATURES Location/Qualifiers
 source
 1..629
 /organism="Bos taurus"
 /mol_type="mRNA"
 /strain="Hereford Shorthorn"
 /db_xref="taxon:9913"
 /clone="CCL003921"
 /sex="female"
 /tissue_type="pooled"
 /dev_stage="Adult"
 /lab_host="XLI-BlueMR" strain"
 /clone_lib="Bos taurus skin cDNA library"
 /note="Organ: skin; Vector: Uni-ZAPXR; Site 1: EcoRI; Site 2: Xho I; Library made from pooled skin of adult female Hereford-Shorthorn."
ORIGIN
 Alignment Scores:
 Pred. No.: 2.73e-28 Length: 629
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x CF768246 (1-629)
 QY 3 GtllleaglnvalilleaaserythrAlaThrglyProgluglnleuThrlenuA 22
 |||||
 Db 21 GAATTCGCCAGGTATTGCTTTACACAGCAACTGGCCTGAGCAGCTTACCTGGCC 80
 QY 23 Proglglnleuileuileuileuileuileuileuileuileuileuileuileu 42
 |||||
 Db 81 CCGGTGAGCTGATCGATCGGAAAAAGAACCGAGTGGATGCTGGAGAGAGCTTA 140
 QY 43 GlnAlaArgGlyLysLysArgGlnlleGlyTrpPheProAlaAsnTyrrValyleu 62
 |||||
 Db 141 CAGGCAAGTGGAGAAAGCCGCAATAGGCTGTTCTGCTAATTCGTAAACTTTTA 200
 QY 63 SerPro 64
 |||||
 Db 201 AGCCCT 206
RESULT 7
LOCUS CD356862
DEFINITION AGENCOURT 14253331 NIH MGC 187 Homo sapiens cDNA clone
 IMAGE:30404497 5', mRNA sequence.
ACCESSION CD356862
VERSION CD356862.1 GI:31128273
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 750)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: NDCM190 row: C column: 02
 High quality sequence stop: 519.
 Location/Qualifiers
 1..750
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30404497"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1b="NIH MGC 187"
 /note="Organ: Blood vessels - aorta, basilar and artery;
 Vector: pDNR-LIB; Site_1: Sfil (ggccatgagcc); Site_2:
 Sfil (ggcgctcgcc); 5' and 3' adaptors were used in
 cloning as follows: 5' adaptor sequence:
 5'-CACGCCCATATGCGC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGCGGCGCCGACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.4 kb
 (range 0.5-4.0 kb). 14/15 colonies contained inserts by
 PCR. This library was enriched for full-length clones and
 was constructed by Clontech Laboratories (Palo Alto, CA).
 Note: this is a NIH MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 3,38e-28 Length: 750
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x CD356862 (1-750)

QY 3 GtialaAGlValIleAlaSerYrTrAlaThGlyProGluGlnLeuThrLeuAla 22
 DB 357 GAATTCGCCAGGTTATTCCTCATACCGCCAGCCGCCGACGACTCTTCGCC 416
 QY 23 ProGlyGlnLeuIleuLeuIleuIleuIleuIleuIleuIleuIleuIleu 42
 DB 417 CCGGTCCAGCTGATTTTGAATCCGAAAAAGAACCCAGGTGGATGTGGGAGAGAGCTG 476
 QY 43 GtialaAGlValIleAlaSerYrTrAlaThGlyProGluGlnLeuThrLeuAla 62
 DB 477 CAAGCAGCGTGGGAAAAAGCGCCAGATAGGCTGTTCAGACTATATATGAAGCTTCTA 536
 QY 63 SerPro 64
 DB 537 AGCCCT 542

RESULT 8
 BQ898718 1062 bp mRNA linear EST 16-AUG-2002
 LOCUS BQ898718
 DEFINITION AGENCOMPT 8118759 lupski_dorsal_root_ganglion Homo sapiens cDNA
 clone IMAGE:6180303 5', mRNA sequence.
 ACCESSION BQ898718
 VERSION BQ898718.1 GI:22290732
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1062)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1AM13562 row: n column: 16
 High quality sequence start: 116
 High quality sequence stop: 760.
 Location/Qualifiers
 1..1062
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6180303"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPOK6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-GTGACCCAGCGTCCG-3' and
 5'-GCTAGTCTAGATCGGAGCGCGCCGCT(15)-3'. Size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

FEATURES

source

Alignment Scores:
 Pred. No.: 5.12e-28 Length: 1062
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: 13 Gaps: 0

ORIGIN

US-09-720-934-2_COPY_999_1062 (1-64) x BQ898718 (1-1062)
 QY 3 GtialaAGlValIleAlaSerYrTrAlaThGlyProGluGlnLeuThrLeuAla 22
 DB 337 GAATTCGCCAGGTTATTCCTCATACCGCCAGCCGCCGACGACTCTTCGCC 396
 QY 23 ProGlyGlnLeuIleuLeuIleuIleuIleuIleuIleuIleuIleu 42
 DB 397 CCGGTCCAGCTGATTTTGAATCCGAAAAAGAACCCAGGTGGATGTGGGAGAGAGCTG 456
 QY 43 GtialaAGlValIleAlaSerYrTrAlaThGlyProGluGlnLeuThrLeuAla 62
 DB 457 CAAGCAGCGTGGGAAAAAGCGCCAGATAGGCTGTTCAGACTATATATGAAGCTTCTA 516
 QY 63 SerPro 64
 DB 517 AGCCCT 522

RESULT 9
 CG508946 446 bp DNA linear GSS 01-OCT-2003
 LOCUS CG508946
 DEFINITION OST59780 Mus musculus 129Sv/Ev Mus musculus genomic clone OST59780,
 genomic survey sequence.
 ACCESSION CG508946
 VERSION CG508946.1 GI:37291689
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	1 (bases 1 to 446)	
AUTHORS	Zambrowicz, B. P., Abuhin, A., Ramirez-Solis, R., Richter, L. J., Plogwitz, U., Beltrandelillo, H., Buxton, E. C., Edwards, J., Pinch, R. A., Frittle, C. J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jia, C., Key, B. W., Jr., Kipp, P., Kohna, B., Ma, Z. Q., Markesich, D., Payne, R., Potter, D. G., Qian, N., Shaw, J., Schrick, J., Shi, Z. Z., Sparks, M. J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A. T.	
TITLE	Wnt1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)	
COMMENT	Contact: Zambrowicz BP OmitBank Lexicon Genetics Incorporated 4000 Research Forest Drive, The Woodlands, TX 77381, USA Email: materials@lexgen.com Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):608-11) Class: Gene trap. Location/Qualifiers 1..446 /organism="Mus musculus" /mol_type="genomic DNA" /strain="129Sv/Ev" /db_xref="taxon:10090" /db_xref="OST59780" /clone="OST59780" /cell_type="embryonic stem cell" /clone_11b="Mus musculus 129Sv/Ev"	
FEATURES		
source		
ORIGIN		
Alignment Scores:		
Score: No.:	6.91e-28	Length: 446
Percent Similarity:	325..00	Matches: 61
Best Local Similarity:	98.39%	Conservative: 0
Query Match:	95.31%	Mismatches: 1
DB:	29	Indels: 0
		Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x CG508946 (1-446)		
QY	3 Glnlleaaglnvalleialasertyrthralathrglyprogluglnleuthrleuala 22	
Db	6 GAAATGGCCAGGTTATTGCTTCTACAGCTGCTACTGCTCCGAAACACTGACCTGACT 65	
QY	23 Progluglnleuileuileuileaglylsyrsanp-progluglytrprrprrgluglglu 42	
Db	66 CCTGGGCGAGCTGATTCGATCCGGAAAAAGAACCCAGGTGATGTGGGAAGGAAACTG 125	
QY	43 GlnlaAATGgLyLsYlSvArGlnlleglytrpPheProAlaAsnTYrVallylsleu 62	
Db	126 CAACTCGAGGGAAGAAACGCCCAATAGCGTGTTTCCAGCAATATATCTCAACTTTTA 185	
QY	63 SerPro 64	
Db	186 AGCCGCC 191	
RESULT 10		
LOCUS	AA063751	537 bp mRNA linear EST 24-SEP-1996
DEFINITION	nt979g10.r1 Soares mouse brain19.5 Mus musculus cDNA clone IMAGE:4483370 5' similar to WP:CI4F5.5 CE01784 SEM-5: C. ELEGANS SEM-5: , mRNA sequence.	
ACCESSION	AA063751	
VERSION	AA063751.1 GI:1557790	
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 537)	
AUTHORS	Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lasky, M., Le, M., Martin, U., Morris, M., Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,	

TITLE
JOURNAL COMMENT

Thelebing,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterson,R.
The Washu-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
Washu-HHMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LBL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:293114
Possible reversed clone: similarity on wrong strand
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 364.

FEATURES
Source
Location/Qualifiers

```

1..537
    /organism="Mus musculus"
    /mol_type="mRNA"
    /db_xref="taxon:10090"
    /clone_id="IMAGE:482370"
    /ded_stage="19.5 dpc total fetus"
    /lab_host="DH10B (ampicillin resistant)"
    /note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dAT) primer [5', TGTTACCACTCGAAGGAGGGCCGCGCATTTTTTTTTTTTTTTT 3'] , double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."
```

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best local Similarity:	Query Match:
8.63e-28	325.00	98.39%	98.39%	95.31%
Length: 537	Matches: 61	Conservative: 0	Mismatches: 1	Indels: 0
Gaps: 0				

US-09-720-934-2_COPY_999_1062 (1-64) x AA063751 (1-537)

QY 3 GlnllelaaglnvallealaserTyThAlathrglyPrroguglnleuThLeuAla 22

Db 175 GAATTGCCAGGTATTATTCCTTACTACGCCTACTGTCCGACAACATCACCCTGGCT 234

QY 23 ProGlyglnleuleulleuileuileargylslasandProsglyglytyrptpgunglygluleu 42

Db 235 CCTGGCAGCTGATTCATTCGATCCGAAAAAGAACCAAGTGATGTGGGAAGAAGAACTG 294

QY 43 GlmlalaarglylslysarsrglnllgAlTyrpheProhlaaseryVallysleuLeu 62

Db 295 CMACTCGAGGAAAAGCCGCGCATAGGGTGGTTCAGCAAAATTAATGCAAACTTCTA 354

QY 63 SerPro 64

Db 355 AGCCCC 360

RESULT 11

A1594919 603 bp mRNA linear EST 15-MAR-2000

LOCUS A1594919

DEFINITION m179g910.y1 Soares mouse b1NM919.5 Mus musculus cDNA clone IMAGE:482370 5' similar to TR:O42287 O42287 INTERSECTION.; mRNA sequence.

ACCESSION A1594919

VERSION A1594919.1 GI:4603967

KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 603)
Marras, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)

TITLE
JOURNAL Contact: Marras M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu

COMMENT This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. This read is a RESEQUENCE of a previously sequenced mouse clone. This read has been verified (found to hit its original self in the correct orientation)
MGI:293114
Seq primer: -40RP from Gibco
High quality sequence stop: 464
POLYA=No.

FEATURES
source
1. 603
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:482370"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NMF19.5"
/note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGATGAGCGGCGCGCATTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Felima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

ORIGIN
Alignment Scores:
Pred. No.: 9.92e-28 Length: 603
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x A1594919 (1-603)

QY 3 GtatttAaGlnValllleAlasertYrThraAlaThGlyProGluGlnleuThrlauA 22
Db 175 GAATATGCCAGGTTATGCTTCTTAAGCTGCTACTGAGGTCCGAAACAACCTCCGGCT 234
QY 23 ProGlyGlnleuileuileuileuileuileuileuileuileuileuileuileu 42
Db 235 CCGGCGAGCTGATCTGATCCGAAAAAGAACCCAGGTGATGCGGGAAGAGAACTG 294
QY 43 GlnAlaArgGly 62
Db 295 CAAGCTCGAGGAGAAAAAGCGCGAGATAGGGTGGTTCCAGCAATATATGCAAACTCTTA 354
QY 63 SerPro 64

Db 355 AGCCCC 360
|||||
RESULT 12
BQ443095 610 bp mRNA linear EST 29-MAY-2002
LOCUS UI-M-EVO-bxf-3-12-0-UI.r1 NIH_BMAP_EVO Mus musculus cDNA clone
DEFINITION IMAGE:5707499 5', mRNA sequence.
ACCESSION BQ443095
VERSION BQ443095.1 GI:21246207
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 610)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
source
1. 610
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5707499"
/issue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP EVO"
/note="Organ: brain; Vector: pYX-Ase; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Ase vector. The library tag sequence located between the Not I site and the polyA tail, is GTGCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hewin Chin, Ph.D., program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 1.01e-27 Length: 610
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x BQ443095 (1-610)

QY 3 GtatttAaGlnValllleAlasertYrThraAlaThGlyProGluGlnleuThrlauA 22
Db 310 GAATATGCCAGGTTATGCTTCTTAAGCTGCTACTGAGGTCCGAAACAACCTCCGGCT 369

Qy	22	ProGlyGluIleuIleuIleuLeaTgLySySaenProGlyGTPTpGUgUglyIleu	42
Db	370	CGTGGGCACTGATTCTGTATCTCGAGAAAAAGAACCCAGGTGATGGTGGGAAGAACA	429
Qy	43	GlnAlaArgGlyLysLysAerGlnIleGlyTTPhePheProAlaAenTYValVysIleu	62
Db	430	CAAGCTTCGAGGGAAAAAGGCCAGATAGGGTGGTTTCAGCAAAATATGTCAAACCTTCTA	489
Qy	63	SerPro 64	
Db	490	AGCCCC 495	
RESULT 13			
BUD054334		694 bp	mRNA linear EST 26-AUG-2000
LOCUS	UI-M-PD0-zbh-n-08-0-UI.r1 NIH BMAP_PD0 Mus musculus cDNA clone		
DEFINITION	IMAGE:6403783 5', mRNA sequence.		
VERSION	BUD054334.1 GI:22494411		
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 694)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Dr. James Lin, University of Iowa CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.jnl.gov This clone was contributed by the Brain Molecular Anatomy Project		

ORIGIN	
Alignment Scores:	
Pred. No.:	1.17e-27
Score:	325.00
Percent Similarity:	98.39%
Best Local Similarity:	98.39%
	length: 694
	Matches: 61
	Mismatches: 1
	Mismatches: 0

Query Match:	95.31%	Indels:	0
DB:	13	Gaps:	0
US-09-720-934-2_COPY_999_1062 (1-64) x BU054334 (1-694)			
Qy	3	GlnIleAaGlnValIleIleAaSerTyrThrAlaThrGlyProGluGlnIleuThreuAla	22
Db	120	GAATATCCCGCAGGATATATGCTCTCCACGCTCTACTGCTGCCGAACAACACCTGAGCT	179
Qy	23	ProGlyGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu	42
Db	180	CTCGGAGAGCTGATTCGTGATCCGAAAAAAGAACCCAGGTGATGTGGGAAGAGAACTG	239
Qy	43	GlnAlaArgGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu	62
Db	240	CAAGCTCGAGGAAAAAAGCGCCAGATAGGGGTGTTCCAGCAAAATTATGTCAAACTTCTA	299
Qy	63	SerPro 64	
Db	300	AGCCCC 305	
RESULT 14			
BM950476			
LOCUS			
DEFINITION	732 bp	mRNA	linear EST 14-MAR-2002
IMAGE:5687260 5', mRNA sequence.			
ACCESSION	BM950476		
VERSION	BM950476.1		
KEYWORDS	GI:19434066		
SOURCE	EST.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryotic Metazoa: Chordata, Craniata, Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 732)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		

```

FEATURES
    source
        Seq primer: PYX-5.
        Location/Qualifiers
            1..732
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6"
                /db_xref="taxon:10090"
                /clone="IMAGE:5687260"
                /tissue_type="whole brain"
                /dev_stage="embryo 18.5 dpc"
                /lab_host="DH10B (T1 phage resistant)"
                /clone_lib="NIH BMAP EHDP"
                /note="Organ: brain; Vector: PYX-ASC; Site_1: EcoR I;
                Site_2: Not I; The library was constructed according to
                Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
                1996. Denatured mRNA was size fractionated on a 1% agarose
                gel. First strand cDNA synthesis was primed with an
                oligo-dT primer containing a Not I site. Double stranded
                cDNA was size selected according to mRNA size fraction,
                ligated with EcoR I adaptor, digested with Not I, and then
                cloned directionally into PYX-ASC vector. The library tag
                sequence located between the Not I site and the polyA
                tail, is CAGCAGCAGC. This library was created for the
                University of Iowa Mouse Brain Molecular Anatomy Project
                (BMAP): 'Gene Discovery in the Developing Mouse Nervous

```

ORIGIN

System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

Alignment Scores:

Pred. No.:	1.25e-27	Length:	732
Score:	325.00	Matches:	61
Percent Similarity:	98.39%	Conservative:	0
Best Local Similarity:	98.39%	Mismatches:	1
Query Match:	95.31%	Indels:	0
DB:	12	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x BM950476 (1-732)

```
QY 3 GtUtlAAGlAgtAallleAlaSerTyrrThAlaThGlyProgluGlnleuThleuAla 22
Db 395 GAAATTGCCAGGTATTTGCTTCTTCCAGCTGCTACTGCTCCGAAACAACCTCAGCTGCT 454
QY 23 ProglYglnleuileuileuileuileuileuileuileuileuileuileuileu 42
Db 455 CTTGGGCACTGATTTCTGATCCGAAAAAGAACCCAGGTGATGCTGGAGAGAACTG 514
QY 43 GlnAlaArgGlyLysLysArgGlnilegLYTrrPheProAlaAsnTyrrAllyleu 62
Db 515 CAAGCTCGAGGAAAAAGCCAGATAGGCTGCTTCCAGCAATTATGCAAACTTCTA 574
QY 63 SerPro 64
Db 575 AGCCCC 580
```

RESULT 15
LOCUS BU614534 782 bp mRNA linear EST 20-FEB-2003
DEFINITION UI-M-EVO-cbg-1-22-0-UI.x1 NIH_BMAP_EVO Mus musculus cDNA clone
ACCESSION BU614534
VERSION BU614534.1 GI:23280749
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 782)
NIH-MGC <http://img.ncbi.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
Location/Qualifiers
1..782
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-EVO-cbg-1-22-0-UI"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP EVO"
/note="Organ: brain; Vector: pYX-Aec; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an

ORIGIN

Alignment Scores:

Pred. No.:	1.35e-27	Length:	782
Score:	325.00	Matches:	61
Percent Similarity:	98.39%	Conservative:	0
Best Local Similarity:	98.39%	Mismatches:	1
Query Match:	95.31%	Indels:	0
DB:	13	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x BU614534 (1-782)

```
QY 3 GtUtlAAGlAgtAallleAlaSerTyrrThAlaThGlyProgluGlnleuThleuAla 22
Db 190 GAAATTGCCAGGTATTTGCTTCTTCCAGCTGCTACTGCTCCGAAACAACCTCAGCTGCT 249
QY 23 ProglYglnleuileuileuileuileuileuileuileuileuileuileuileu 42
Db 250 CTTGGGCACTGATTTCTGATCCGAAAAAGAACCCAGGTGATGCTGGAGAGAACTG 309
QY 43 GlnAlaArgGlyLysLysArgGlnilegLYTrrPheProAlaAsnTyrrAllyleu 62
Db 310 CAAGCTCGAGGAAAAAGCCAGATAGGCTGCTTCCAGCAATTATGCAAACTTCTA 369
QY 63 SerPro 64
Db 370 AGCCCC 375
```

Search completed: July 1, 2004, 19:48:54
Job time : 1931.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 1345.81 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_1080_1138
Sequence: 1 AYCQYIGMIDYTRQNDDELA.....GEVNGVGLFPBNYVLTLD 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q=/cg2 1/USPTO_fastap.p/US09720934/runat 30062004_064539_13430/app.query.fasta_1.1386
-DB=GenEmbl -QFMT=step -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-INITs-blts -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAVYSIZE=500 -MINLEN=2000000000
-USER=US09720934.@CGN 1.1 7509 @runat 30062004_064539_13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_sts:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_nam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	323	100.0	1996	9	AF180522	AF180522 Homo sapi
2	323	100.0	3231	6	BD205037	BD205037 Isolated
3	323	100.0	3241	9	HSU61166	U61166 Human SH3 d
4	323	100.0	4321	9	BSM806384	BXS38175 Homo sapi
5	323	100.0	5195	6	BD205035	BD205035 Isolated
6	323	100.0	5199	6	BD205033	BD205033 Isolated
7	323	100.0	5287	9	AF064243	AF064243 Homo sapi
8	323	100.0	5381	9	AF114488	AF114488 Homo sapi
9	323	100.0	5458	6	BD205034	BD205034 Isolated
10	323	100.0	6439	9	AF114487	AF114487 Homo sapi
11	323	100.0	7247	9	AF064244	AF064244 Homo sapi
12	323	100.0	12015	9	AP000312	AP000312 Homo sapi
13	323	100.0	100000	9	AP000193	AP000193 Homo sapi
14	323	100.0	114929	9	AP000050	AP000050 Homo sapi
15	323	100.0	151516	9	AP000117	AP000117 Homo sapi
16	323	100.0	171361	9	BS000184	BS000184 Pan trogl
17	323	100.0	184526	2	AC144367	AC144367 Papio anu
18	323	100.0	194157	2	AC144368	AC144368 Papio anu
19	323	100.0	340000	9	AF001718	AF001718 Homo sapi
20	322	99.7	78190	9	AC106760	AC106760 Homo sapi
21	322	99.7	179264	2	AC092312	AC092312 Homo sapi
22	319	98.8	203423	2	AC137544	AC137544 Gallus ga
23	317	98.1	142974	2	AC129075	AC129075 Felis cat
24	317	98.1	157107	2	AC140666	AC140666 Sus scrof
25	317	98.1	169774	2	AC138787	AC138787 Sus scrof
26	313	96.9	1133	10	AF169621	AF169621 Mus muscu
27	313	96.9	3723	10	AF132478	AF132478 Mus muscu
28	313	96.9	5145	10	AF132481	AF132481 Mus muscu
29	313	96.9	217131	10	AC126053	AC126053 Mus muscu
30	312	96.6	224721	2	AC139632	AC139632 Bos tauri
31	308	95.4	3812	10	AF132672	AF132672 Rattus no
32	308	95.4	4025	10	AF127798	AF127798 Rattus no
33	308	95.4	135924	2	AC115529	AC115529 Rattus no
34	307	95.0	178801	2	AC121075	AC121075 Canis fam
35	307	95.0	235426	2	AC121074	AC121074 Canis fam
36	296	91.6	76179	2	AL606725	AL606725 Dario rer
37	296	91.6	104334	5	AL606751	AL606751 Zebrafish
38	296	91.6	163197	5	BX005416	BX005416 Zebrafish
39	296	91.6	182740	2	AC138436	AC138436 Dario rer
40	296	91.6	225535	2	BX470235	BX470235 Dario rer
41	295	91.3	4103	5	AF032118	AF032118 Xenopus l
42	282	87.3	113146	2	AC138440	AC138440 Tetradon
43	282	87.3	127051	2	AC138485	AC138485 Tetradon
44	282	87.3	134408	2	AC139628	AC139628 Takifugu
45	282	87.3	136937	2	AC139627	AC139627 Takifugu

RESULT 1

ALIGNMENTS

AF180522 AF180522 1996 bp mRNA linear PRI 05-SEP-1999
LOCUS Homo sapiens interseclin short form 2 (ITSN) mRNA, partial cds.
DEFINITION AF180522
ACCESSION AF180522.1 GI:5823551
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1996)
Tsuda, L.O., Kvasna, S.M., Skripkina, I.Y., Anoprienko, O.V.,
Slavov, D., Tassone, F., Rynditch, A.V. and Gardiner, K.
TITLE Mouse homologs of human chromosome 21 genes
JOURNAL Unpublished
2 (bases 1 to 1996)
Tsuda, L.O., Kvasna, S.M., Skripkina, I.Y., Anoprienko, O.V.,
Slavov, D., Tassone, F., Rynditch, A.V. and Gardiner, K.
REFERENCE
AUTHORS Direct Submission
Submitted (25-AUG-1999) Department of Molecular Oncogenetics,
Institute of Molecular Biology and Genetics of National Academy of
Sciences of Ukraine, 150, Zabolotnogo str., Kiev 252627, Ukraine
TITLE
JOURNAL Location/Qualifiers
FEATURES
Source 1..1996
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="Brain"
/dev_stage="infant"
<1..1996
/gene="ITSN"
<1..566
/gene="ITSN"
/codon_start=3
/product="interseclin short form 2"
/protein_id="AAB53183.1"
/db_xref="GI:5823552"
/translation="KSYVKILSGIPRKSTSDSSSPASLKVSPAPKPVSGE
IAQYIASYTAAGPEQTLAPQQLIRKKNRGMWEGELQKRGKKRQIEMFPANVYKL
LSPEKSKITPEPRKSTALAAVCQVIGMDYDTAQNDELAFNKQIINVINKEDPDWM
KGEVNGVGVPSPSNVYKLTMDPSQQ"
ORIGIN
Alignment Scores:
Pred. No.: 4.01e-32 Length: 1996
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_1080_1138 (1-59) x AF180522 (1-1996)
QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
Db 369 GCACTGTGCCAGCGATTGGAGTGAACACTACACCGCGCAAGATGACGATGAGCTGGCC 428
QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40
Db 429 TTCAACAAGGGCCGATCATCAACGCTCTCAACAAGGAGGACCTGACGTGTGGAAGA 488
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
Db 489 GAAGTCATGACACAGTGGGGCTCTCCCATCCCAATTATGTGAAGCTGACCAAGAC 545
RESULT 2 BD205037 8231 bp DNA linear PAT 17-JUL-2003
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205037

VERSION BD205037.1 GI:33014807
KEYWORDS JP 2002511267-A/5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 3231)
Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 5 16-APR-2002;
CEDARS SIVAI HEALTH SYSTEM ET AL
COMMENT
OS Homo sapiens (human)
PN JP 2002511267-A/5
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R KORENBERG, XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C1201/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1..3231
/organism="Homo sapiens (human)"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.68e-32 Length: 3231
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_1080_1138 (1-59) x BD205037 (1-3231)
QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
Db 1730 GCACTGTGCCAGCGATTGGAGTGAACACTACACCGCGCAAGATGACGATGAGCTGGCC 1789
QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40
Db 1790 TTCAACAAGGGCCGATCATCAACGCTCTCAACAAGGAGGACCTGACGTGTGGAAGA 1849
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
Db 1850 GAAGTCATGACACAGTGGGGCTCTCCCATCCCAATTATGTGAAGCTGACCAAGAC 1906
RESULT 3 HSU61166 3241 bp mRNA linear PRI 23-JUL-1996
LOCUS Human SH3 domain-containing protein SH3P17 mRNA, complete cds.
DEFINITION U61166
ACCESSION U61166.1 GI:1438932
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 3241)
Sparks, A.B., Hoffman, N.G., McConnell, S.J., Fowlkes, D.M. and
Kay, B.K.
TITLE Cloning of ligand targets: systematic isolation of SH3
domain-containing proteins

VERSION BD205035.1 GI:33014805
 KEYWORDS JP 2002511267-A/3.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 5195)
 Korenberg, J.R. and Chen, X.N.
 Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
 Patent: JP 2002511267-A 3 16-APR-2002;
 CEDARS SINAI HEALTH SYSTEM ET AL
 COMMENT OS Homo sapiens (human)
 PN JP 2002511267-A/3
 PD 16-APR-2002
 PF 16-APR-1999 JP 2000543610
 PR 16-APR-1998 US 60/082007
 PI JULIE R KORENBERG, XIAO NING CHEN
 PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
 PC C12Q1/68,
 PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
 CC Isolated SH3 gene relating to myeloproliferative disorders and leukemia
 CC and utilization thereof.
 FH Key Location/Qualifiers
 FT source 1..5195
 FT Location/Qualifiers
 1..5195
 /organism="Homo sapiens (human)"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1..1e-31 Length: 5195
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_1080_1138 (1-59) x BD205035 (1-5195)
 QY 1 A1aValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
 Db 3689 GCAGTGGCCAGCGATGGATGGAGTACAGACTACACCGCGAGAATGACGATGAGCTGGCC 3748
 QY 21 PheAnlySGlyGlnIleIleAsnValLeuAnlySGluAspProAspTyrTrpLysGly 40
 Db 3749 TTCAACAAGAGGCGCATCATCATCGCTCTCAACAAGAGGAGACCTGACTGGTGGAAAGA 3808
 QY 41 GluValAenGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThraSp 59
 Db 3809 GAAGTCATATGACAAAGTGGGGCTCTCCATCCCAATTATGTGAAGCTGACCAAGAC 3865
 RESULT 6 BD205033 5199 bp DNA linear PAT 17-JUL-2003
 LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
 DEFINITION leukemia and utilization thereof.
 ACCESSION BD205033
 VERSION BD205033.1 GI:33014803
 KEYWORDS JP 2002511267-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 5199)
 Korenberg, J.R. and Chen, X.N.
 Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof

JOURNAL Patent: JP 2002511267-A 1 16-APR-2002;
 CEDARS SINAI HEALTH SYSTEM ET AL
 COMMENT OS Homo sapiens (human)
 PN JP 2002511267-A/1
 PD 16-APR-2002
 PF 16-APR-1999 JP 2000543610
 PR 16-APR-1998 US 60/082007
 PI JULIE R KORENBERG, XIAO NING CHEN
 PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
 PC C12Q1/68,
 PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
 CC Isolated SH3 gene relating to myeloproliferative disorders and leukemia
 CC and utilization thereof.
 FH Key Location/Qualifiers
 FT source 1..5199
 FT Location/Qualifiers
 1..5199
 /organism="Homo sapiens (human)"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1..11e-31 Length: 5199
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_1080_1138 (1-59) x BD205033 (1-5199)
 QY 1 A1aValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
 Db 3445 GCAGTGGCCAGCGATGGATGGAGTACAGACTACACCGCGAGAATGACGATGAGCTGGCC 3504
 QY 21 PheAnlySGlyGlnIleIleAsnValLeuAnlySGluAspProAspTyrTrpLysGly 40
 Db 3505 TTCAACAAGAGGCGCATCATCATCGCTCTCAACAAGAGGAGACCTGACTGGTGGAAAGA 3564
 QY 41 GluValAenGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThraSp 59
 Db 3565 GAAGTCATATGACAAAGTGGGGCTCTCCATCCCAATTATGTGAAGCTGACCAAGAC 3621
 RESULT 7 AF064243 5287 bp mRNA linear PRI 21-NOV-1998
 LOCUS Homo sapiens interectin short form mRNA, complete cds.
 DEFINITION AF064243
 ACCESSION AF064243
 VERSION AF064243.1 GI:3859852
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 5287)
 Guipponi, M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and Antonarakis, S.E.
 Two isoforms of a human interectin (ITSN) protein are produced by brain-specific alternative splicing in a stop codon
 JOURNAL Genomics 53 (3), 369-376 (1998)
 MEDLINE 99017974
 PUBMED 9799604
 REFERENCE 2 (bases 1 to 5287)
 AUTHORS Guipponi, M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and Antonarakis, S.E.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue Michel-Servet, Geneva 4 CH-1211, Switzerland
 FEATURES Location/Qualifiers

US-09-720-934-2_COPY_1080_1138 (1-59) x AF114488 (1-5381)

QY 1 AlAlaValGysGInValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGlnLeuAla 20
DB 3734 GAGGTGCGCAGGATGGATGGATGACATCAACCGCGCAGAAAGACGATGAGCTGGCC 3793

QY 21 PheAsnLysGlyGlnIleIleAsnValIleuAsnLysGlyAspProAspTyrTrpLysGly 40
DB 3794 TTCAACAAGGGCCAGATCATCAACGCTCTCAACAAGAGAGACCCTGATCGTGGAAAGGA 3853

QY 41 GluValaGngLysGlnValaGlyLeuPheProSerAsnTyrValLysLeuThrAsp 59
DB 3854 GAAGTCATGACAGACAGTGGGGCTCTCCCATCCCAATTATGTGAAGCTGACCAAGAC 3910

RESULT 9
BD205034 5458 bp DNA linear PAT 17-JUL-2003
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205034
VERSION BD205034.1 GI:33014804
KEYWORDS JP 2002511267-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Korenberg, J.R. and Chen, X.N.
AUTHORS Isolated SH3 gene relating to myeloproliferative disorders and
TITLE leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 2 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL
COMMENT OS Homo sapiens (human)
PN JP 2002511267-A/2
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R KORENBERG, XIAO NING CHEN
PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
PC C1201/68,
PC G01N33/68//A61K46/00, C12N15/00, C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

CC leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1.5458
FT /organism="Homo sapiens (human)".
FT Location/Qualifiers
1.5458
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1.16e-31
Score: 323.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BD205034 (1-5458)

QY 1 AlAlaValGysGInValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGlnLeuAla 20
DB 3732 GCAATGGCCAGGATGGATGGATGACATCAACCGCGCAGATGACGATGAGCTGGCC 3791

QY 21 PheAsnLysGlyGlnIleIleAsnValIleuAsnLysGlyAspProAspTyrTrpLysGly 40
DB 3792 TTCAACAAGGGCCAGATCATCAACGCTCTCAACAAGAGAGACCCTGATCGTGGAAAGGA 3851

QY 41 GluValaGngLysGlnValaGlyLeuPheProSerAsnTyrValLysLeuThrAsp 59
DB 3852 GAAGTCATGACAGACAGTGGGGCTCTCCCATCCCAATTATGTGAAGCTGACCAAGAC 3908

RESULT 10
AF114487 6439 bp mRNA linear PRI 16-JUL-2002
LOCUS Homo sapiens intersecin long isoform (ITSN) mRNA, complete cds.
DEFINITION AF114487
ACCESSION AF114487.1 GI:4808822
VERSION AF114487.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 6439)
AUTHORS Pucharcos, C., Fuentes, J.J., Casas, C., de la Luna, S., Alcantara, S.,
Arbones, M.L., Soriano, E., Estivill, X., and Pritchard, M.
TITLE Alu-splice cloning of human intersecin (ITSN), a putative
multivalent binding protein expressed in proliferating and
differentiating neurons and overexpressed in Down syndrome
JOURNAL Eur. J. Hum. Genet. 7 (6), 704-712 (1999)
MEDLINE 99415290
PUBMED 10482960
REFERENCE 2 (bases 1 to 6439)
AUTHORS Pucharcos, C., Fuentes, J.J., Pritchard, M., and Estivill, X.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
Cancer Research Institute, L'Hospitalet de Llo., Avia,
Castelldefels km. 2.7, Barcelona 08907, Spain
FEATURES
source
1.6439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1.6439
/gene="ITSN"
269.5434
/gene="ITSN"
/codon_start=1
/product="intersecin long isoform"
/protein_id="AAD2952.1"
/db_xref="GI:4808823"
/translation="MAQPTTPFGSGSLDIMAITYERAKHQDQPSLRPISGFTGDOA
RNFPPQSLPQVLAQIWLADNMNDQMDQVRSIAKLILKLQGLQSLPAPVPM
KQDFVAISSAPAFMGQIASNPPLTAVAPVPMGSIPIVGMSPTIASVPTAAVPLAN
GAPVLIQPLPAFAIPAATLIPKSSFSRSGPSQNTLQKASQFDVASVPTAAVAVP
QSSRLKYROLFNSHDKTMSGHLTGPOARTILMOSSLPQOALASITNLSIDIDQKLT
BEFTLAHMLIDVAMSGQPLPVLPEVYIPSEFRVRSQSGISVISTSVQDRLPEPV
LEDECOOLEKTLPTTPEDKRENERNERNELERKQALIBQOREGRIALDEABQ
RKERQROERKQLELEKQERELRQREERERKTEREAKRELERQRELE
RNRQRLNQNKEQEDIVLAKRKTLTELELAKNDKQLEKLDIDRLTTRQ
EIESTNSRELRLAETHLOQLOESQOMLERLIPKQILNDQLOVOYQNSLHSDIV
TLKALAKELARQHLRDQLEVEKTRSKQLEIDFNNQKLEIRLHNKQLOKOS
MEARLKQKQERKILEKQKEARQKQLEKQKQLEKQKQLEKQKQLEKQKQLEKQKQ
KRESYVKKQ
KUYVYALYPRSSSHDEITQPDLYMKGWYDEQSGEPHGLGELKKTQMPFA
NYAKTIPENVPAPKPVITDSTSAFPAKLAIRETPALVATSSPSTTPNNWADPST
WPTSTNEKPEPDWDAMAAQPSLTPVPSAGLQKRSATPAPATSSSPSVYQGEKVE
GLQALALYPMWAKKDNHINFNKNDVITVLEQDMMWGEVQGGPKSVYKLSGP
IRKSTMSDSSSESSEPSALIKRVASPAKPVVSGEFTIMYTESSEQGLTFQOGDVL
VTKKDGMWMTGIVDGKAGVPPSNVYRLKDSGSGTAGKTSIGKKEPILAOVITATYT
GPBQTLTAPQULIRKKNQGGEGELQARKKQROGMPANYVYKLSFETSLTTP
EPKSTALAAVCQVITGYDTTAQNDDELAERKQIINVLNKEDDQWKGEGNQGVL
PSNYVKLTTMDPSQWQCSDLHLMITPERKQGYIHELIVTEENVVNDLQVTEI
FQKLMSEELLTEKEVAMIIVNWKELITMCNLIKILKALRVKKSQSGEPMVQIDILS
AOLPMDQVYRPFCSROLNGAALIOCKTDEAPDEKFEVYKRLMDPRCKGMPSSPTLK
MORTVRYPLIINKILENTPENHPDHSILKHALKEAEELGQVNSGVREKENSDELEW
QAHVQCGELSEQVYFNSVTNCLGRKRLHSGKLYKAKSNKEIYGFLENDPLILQITK
PLSSGTDKVFSPKSNIQYRMVKTPIPLNEVLVLPDPSGDEBTPHSHIDRYATLR
AESINERTAWQKIKASLELYIETEKKKRERKAYLVRSQATIGRLVWNVVEGIELPK

ORIGIN

CRSHGKSNPYCEVTWMSQCHITKTIDTILNPKMNSCOFFIRDLBOELCTIVFERDQ
FSPDDELGRTEIRIVADIKKDQSGKGVTKCLLHVEVPTGEIVRDLQLDEDE"

Alignment Scores:

Pred. No.:	1.39e-31	Length:	6439
Score:	323.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x AF114467 (1-6439)

QY 1 AlAlaValCyGslGlnVal11leg1MeTyTyrAspTyThrAlaGlnAsnAspAspGluLeuAla 20
Db 3734 GCAGTGTGCCAGGTGATGATGGATGTACGACTACACCCGCGAGATGACGATGAGCTGGGCC 3793
QY 21 PheAsnLysGlyGlnIleIleAsnVal1LeuAsnLysGluAspProAspTyrTyrPlyGly 40
Db 3794 TTCAACAAGGCGCCAGATCATCAACGCTCTCAACAAGAGGACCTGACTGTGGAAGA 3853
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyTyrAllylsLeuThrThrAsp 59
Db 3854 GAAGTCAATGACCAAGTGGGGCTCTTCCCATCCATATATGTGAAGCTGACCAAGAC 3910

RESULT 11
AF064244 7247 bp mRNA linear PRI 21-NOV-1998
LOCUS AF064244 Homo sapiens intersecin long form mRNA, complete cds.

DEFINITION AF064244
ACCESSION AF064244
VERSION AF064244.1 GI:3859854
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 7247)
AUTHORS Guiponi, M., Scott, H. S., Chen, H., Schebesta, A., Rossier, C. and Antonarakis, S. E.
TITLE Two isoforms of a human intersecin (ITSN) protein are produced by brain-specific alternative splicing in a stop codon
GENOMICS 53 (3), 369-376 (1998)
JOURNAL MEDLINE 99017974
PUBMED 9799604

REFERENCE 2 (bases 1 to 7247)
AUTHORS Guiponi, M., Scott, H. S., Chen, H., Schebesta, A., Rossier, C. and Antonarakis, S. E.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue Michel-Servet, Geneva 4 CH-1211, Switzerland
FEATURES
source Location/Qualifiers
1..7247
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="fetus"
107..5272
/codon_start=1
/product="intersecin long form"
/protein_id="AAC78611.1"
/db_xref="GI:3859854"
/translation="MAQPTPPFGSLDIWATVEERAKHDQPHSKPISGFIITGQA
RNFPOSGLPOPLAIOIWLADNNDGMOVQFSLAMKLIKLOGLPSALPEVM
KOOPVAISSAPRPGMGIAEMPLTAVPMPGSIIVGMSPIVSSVPTAAPPAN
GAPVIOPLPAPAPAPATAPKSSPSRSGSGLNKLQKAGSFVAPVAPKAVP
QSRKLRQLPNSHDKTMSHLTGPQKRTILMSSLPQQLAIKMLSDIDQGLTA
EERTLAMHLIDVAMSGQPLPVPILPEYIPSPFRVSSGSIIVTSISVQVRPEEPV
LEDEQQLERKLPVTEDEKKRENERGNTLEKRRQALQQRKEQERLALQLRACQ
RKRREROERKRKQLELEKQLEKQRELEQRERREERREKREIRREARKELEKQLEME

CDS

US-09-720-934-2_COPY_1080_1138 (1-59) x AF064244 (1-7247)

ORIGIN

Alignment Scores:
Pred. No.: 1.57e-31
Score: 323.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9

QY 1 AlAlaValCyGslGlnVal11leg1MeTyTyrAspTyThrAlaGlnAsnAspAspGluLeuAla 20
Db 3572 GCAGTGTGCCAGGTGATGATGGATGTACGACTACACCCGCGAGATGACGATGAGCTGGGCC 3631
QY 21 PheAsnLysGlyGlnIleIleAsnVal1LeuAsnLysGluAspProAspTyrTyrPlyGly 40
Db 3632 TTCAACAAGGCGCCAGATCATCAACGCTCTCAACAAGAGGACCTGACTGTGGAAGA 3691
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyTyrAllylsLeuThrThrAsp 59
Db 3692 GAAGTCAATGACCAAGTGGGGCTCTTCCCATCCATATATGTGAAGCTGACCAAGAC 3748

RESULT 12
AF000312 12015 bp DNA linear PRI 24-MAY-2003
LOCUS AF000312 Homo sapiens genomic DNA, chromosome 21 clone:RP1-201f12, complete
DEFINITION sequence.
ACCESSION AF000312
VERSION AF000312.2 GI:31071646

KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (1999)
REFERENCE
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT On May 23, 2003 this sequence version replaced gi:4835681.
FEATURES
Source
Location/Qualifiers
1..12015
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q"
/clone="Rpl-201F12"
ORIGIN
Alignment Scores:
Pred. No.: 2,68e-31 Length: 12015
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_1080_1138 (1-59) x AP0000312 (1-12015)
QY 1 AAlaValCysGlnValIleGlyMetCysAspTyrThrAlaGlnAsnAspAspGlnLeuAla 20
Db 2652 GCAGTGTCCGCGGATGGATGATGACATGACACCGCGCAAAAGACGATGAGCTGCGC 2711
QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGlyAspProAspTyrTrpLysGly 40
Db 2712 TTCAACAAGGGCCACATATATCACTCTCTCAACAAGGAGACCTGACTGGTGGAAAGGA 2771
QY 41 GluValaGnglyGlnValaGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
Db 2772 GAAGTCATGACAAAGTGGGGCTCTCCATCCATTCATTTGTGAAGCTGACCAACAGAC 2828
RESULT 13
AP000193 100000 bp DNA linear PRI 08-JAN-2000
LOCUS Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,
DEFINITION clone 076C10-f32B9, segment 20/21, complete sequence.
ACCESSION AP000193
VERSION AP000193.1 GI:4826584
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y. and Sakaki, Y.
TITLE Homo sapiens 2,083,744bp genomic DNA of 21q22.1 (REGION:
D21S226-AML CLONE RANGE: 076C10-f32B9)
JOURNAL Published Only in Database (1999)
REFERENCE
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y. and Sakaki, Y.

TITLE Direct Submission
JOURNAL Submitted (10-MAY-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail: hattori@gsc.riken.go.jp,
URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,
Fax: 81-42-778-9924)
COMMENT E. coli transposon insertion: The present data does not contain E.
coli transposon sequences which integrated in the
original/previous sequences. We determined the boundary between
the insertion and genomic sequences experimentally, removed the
insertion sequences, reconstituted the present data. The sequencing
project is supported by Japan Science Technology Corporation (JST)
and The Institute of Physical and Chemical Research (RIKEN).
FEATURES
Source
Location/Qualifiers
1..100000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
ORIGIN
Alignment Scores:
Pred. No.: 2,52e-30 Length: 100000
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_1080_1138 (1-59) x AP000193 (1-100000)
QY 1 AAlaValCysGlnValIleGlyMetCysAspTyrThrAlaGlnAsnAspAspGlnLeuAla 20
Db 64215 GCAGTGTCCGCGGATGGATGATGACATGACACCGCGCAAAAGACGATGAGCTGCGC 64274
QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGlyAspProAspTyrTrpLysGly 40
Db 64275 TTCAACAAGGGCCACATATATCACTCTCTCAACAAGGAGACCTGACTGGTGGAAAGGA 64334
QY 41 GluValaGnglyGlnValaGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
Db 64335 GAAGTCATGACAAAGTGGGGCTCTCCATCCATTCATTTGTGAAGCTGACCAACAGAC 64391
RESULT 14
AP000050 114929 bp DNA linear PRI 20-NOV-1999
LOCUS Homo sapiens genomic DNA, chromosome 21q22.1, segment 21/28,
DEFINITION complete sequence.
ACCESSION AP000050
VERSION AP000050.1 GI:3132360
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998)
REFERENCE
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1998) Masahira Hattori, Kitasato University,
Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1,
Sagamihara 228, Japan (E-mail: hattori@hgc.ims.u-tokyo.ac.jp,
Tel: 0427-78-9732, Fax: 0427-78-9561)
COMMENT This sequence is conducted by Kitasato University JST sequencing
Laboratory as a JST sequencing team.
Principal Investigator: Yoshiyuki Sakaki Ph.D.
Phone: +81-3-5449-5622, Fax: +81-3-5449-5445,
sakaki@hgc.ims.u-tokyo.ac.jp

Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D. The sequence is submitted by: Human Genome Sequencing in ALIS project of JST
Japan Science and Technology Corporation (JST)
5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0028 Japan
For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.alis.tokyo.jst.go.jp/HGS/top.html>) or send email to webmaster@www.alis.tokyo.jst.go.jp.

FEATURES

source

Location/Qualifiers
1. 114929
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="130N6-149C3f_2"

ORIGIN

Alignment Scores:

Pred. No.: 2.92e-30 Length: 114929
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AP000050 (1-114929)

QY 1 AAlvalCysGlnValIleGlyMetTYrAspTYrThrAlaGlnAsnAspGluLeuAla 20
27628 GCACTGCGCAGGTGATGGATGATGACTACACCGCGAGATGACGATGAGCTG6CC 27687
QY 21 PheAsnIysGlyGlnIleIleAsnValLeuAsnIysGluAspProAspTPTPTlyGly 40
27688 TTCACACAGAGGCGCAGATCATCAAGCTCTCCACACAGAGAGACCTGACTGCTGGAAGA 27747
QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTYrValIysLeuThrThrAsp 59
27748 GAAGTCATGAGACAGAGGGGCTCTTCCATCCATATTATGTGAAGCTGACACAGAC 27804
RESULT 15
AP000117
LOCUS 151516 bp DNA linear PRI 25-MAY-2002
DEFINITION Homo sapiens genomic DNA of 21q22.1 GART and AML related,
AP000117
ACCESSION Q78C10-149C3 region, segment 20/20.
VERSION AP000117.1 GI:4730851
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens 2,051,516bp genomic DNA of 21q22.1 GART and AML region
JOURNAL Published Only in Database (1999)
REFERENCE
AUTHORS Hirakawa,M., Yamaguchi,H., Imai,K. and Shimada,U.
TITLE Direct Submission
JOURNAL Submitted (15-APR-1999) Mika Hirakawa, Japan Science and Technology Corporation (JST), Advanced Databases Department, 5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081, Japan (E-mail:mika@tokyo.jst.go.jp, URL:<http://www.alis.tokyo.jst.go.jp/>, Tel:81-3-5214-8491, Fax:81-3-5214-8470)
This sequence is conducted by Kitasato University JST sequencing laboratory as a JST sequencing team.
Principal Investigator: Yoshiyuki Sakaki Ph.D.
Phone: +81-3-5449-5622, Fax: +81-3-5449-5445,
sakaki@ngc.ims.u-tokyo.ac.jp
Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D. The sequence is submitted by Human Genome Sequencing in ALIS project of JST.

FEATURES

source

Japan Science and Technology Corporation (JST)
5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081 Japan
For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.alis.tokyo.jst.go.jp/HGS/>) or send email to webmaster@www.alis.tokyo.jst.go.jp.
Location/Qualifiers
1. 151516
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
55614..55714
/note="SHGC-37625;The location is between each flanking site of PCR primers."
/db_xref="GDB:6458640"
65945..66167
/note="Cda0x03;The location is between each flanking site of PCR primers."
/db_xref="GDB:441363"
65993..66263
/note="TIGR-A002U42;The location is between each flanking site of PCR primers."
/db_xref="GDB:4586120"
105562..105722
/standard_name="D21S235"
/note="KM1012/KM914;The location is between each flanking site of PCR primers."
/db_xref="GDB:185299"
105703..105879
/standard_name="D21S235"
/note="KM1014/KM877;The location is between each flanking site of PCR primers."
/db_xref="GDB:185300"
106314..106441
/note="SHGC-51824;The location is between each flanking site of PCR primers."
/db_xref="GDB:6464287"
106333..106592
/standard_name="D21S324"
/note="G52F04;The location is between each flanking site of PCR primers."
/db_xref="GDB:190697"
125278..125448
/note="SHGC-51805;The location is between each flanking site of PCR primers."
/db_xref="GDB:6464259"
125278..125381
/note="SHGC-44617;The location is between each flanking site of PCR primers."
/db_xref="GDB:6459978"
127391..127601
/standard_name="D21S1940"
/note="N1B1972;The location is between each flanking site of PCR primers."
/db_xref="GDB:626346"
136862..137057
/standard_name="D21S1977"
/note="SHGC-2817;The location is between each flanking site of PCR primers."
/db_xref="GDB:681529"
136914..137060
/standard_name="D21S1980"
/note="SHGC-12120;The location is between each flanking site of PCR primers."
/db_xref="GDB:681538"

ORIGIN

Alignment Scores:
Pred. No.: 3.92e-30 Length: 151516
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9

Mismatches: 0
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AP000117 (1-151516)

QY	1	AlaValCysGlnValIleGlyMet	TrpAspTyrThrAlaGlnAsnAspGluLeuAla	20
DB	64215	GCAGTGTCCCGAGTGTGGATGACACTACACCGCGCAGAAATGACGATGAGCTGSCC		64274
QY	21	PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTrpTrpLysGly		40
DB	64275	TTCAACAAGGCGCCAGATCATCAACGCTCTCAACAAGAGAGACCCCTGACTGGTGGAAAGGA		64334
QY	41	GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrasp		59
DB	64335	GAAATCAATGACACAGTGGGCTCTTCCCATCCAAATTATGTGAAGTGTACCAACAGAC		64391

Search completed: July 1, 2004, 16:16:09
Job time : 1391.81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 41.3138 Seconds

(without alignments)
792.522 Million cell updates/sec

Title: US-09-720-934-2_COPY_1080_1138
Sequence: 1 AVOGVIGMYDYAQNDELIA.....GEVNGQVGFPPSNVYKLTID 59

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+pn.model -DBV=xlp
-Q=/cgn2_1/USFTO_spool_p/US09720934/rnuc 30062004_064540_13455/app.query.fasta_1.1386
-DB=Issued Patents NA -QMT=fastp -SUPFIX=rm1 -MINMATCH=0.1 -LOOPT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720934@cgn1_1.284@runcat 30062004_064540_13455 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOBJECT -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	266	82.4	747	4	US-08-630-915A-39
2	266	82.4	2873	4	US-08-630-915A-193
3	160	49.5	1392	1	US-08-475-894-3
4	160	49.5	1392	1	US-08-484-710-3
5	160	49.5	1392	2	US-08-484-710-3
6	160	49.5	1392	2	US-08-484-710-3
7	160	49.5	1392	2	US-08-474-697-3
8	160	49.5	1659	1	US-08-475-894-1
9	160	49.5	1659	1	US-08-484-710-1
10	160	49.5	1659	2	US-08-484-710-1
11	160	49.5	1659	2	US-08-474-697-1
12	156	48.3	1045	1	US-08-671-354-1
					Sequence 5, Appl

13	156	48.3	1045	1	US-08-484-710-5	Sequence 5, Appl
14	156	48.3	1045	2	US-08-484-709-5	Sequence 5, Appl
15	156	48.3	1045	3	US-08-474-697-5	Sequence 5, Appl
16	154	47.7	1430	2	US-08-549-004A-15	Sequence 15, Appl
17	154	47.7	1430	3	US-09-051-982A-15	Sequence 15, Appl
18	152	47.1	813	4	US-09-023-655-1238	Sequence 138, Ap
19	152	47.1	933	4	US-09-023-655-1238	Sequence 138, Ap
20	152	47.1	933	2	US-08-641-640-1	Sequence 1, Appl
21	152	47.1	984	4	US-09-765-298A-5	Sequence 1, Appl
22	152	47.1	1072	1	US-07-906-349A-2	Sequence 5, Appl
23	152	47.1	1072	1	US-08-167-035-5	Sequence 5, Appl
24	152	47.1	1072	1	US-08-167-035-5	Sequence 5, Appl
25	152	47.1	1072	1	US-08-167-035-5	Sequence 5, Appl
26	152	47.1	1072	2	US-08-208-887A-5	Sequence 49, Appl
27	152	47.1	1072	2	US-08-539-005-49	Sequence 5, Appl
28	152	47.1	1072	2	US-08-539-005-49	Sequence 5, Appl
29	152	47.1	1109	3	US-08-664-962B-5	Sequence 2, Appl
30	152	47.1	1109	3	US-09-311-743-5	Sequence 2, Appl
31	150	46.4	1661	2	US-08-815-176-2	Sequence 5, Appl
32	150	46.4	1661	2	US-08-815-176-2	Sequence 2, Appl
33	148	45.8	1710	4	US-08-630-915A-189	Sequence 189, App
34	144	44.6	1640	4	US-09-023-655-650	Sequence 650, App
35	143.5	44.4	1347	4	US-09-833-381-1937	Sequence 931, App
36	142	44.0	874	4	US-08-942-423-68	Sequence 91, Appl
37	142	44.0	1458	2	US-08-942-423-68	Sequence 68, Appl
38	142	44.0	2003	4	US-08-942-423-1	Sequence 25, Appl
39	142	44.0	2003	4	US-08-630-915A-25	Sequence 32, Appl
40	141	43.7	949	1	US-08-167-035-32	Sequence 50, Appl
41	141	43.7	949	1	US-08-167-035-50	Sequence 32, Appl
42	141	43.7	949	1	US-08-208-887A-32	Sequence 32, Appl
43	141	43.7	949	2	US-08-539-005-50	Sequence 50, Appl
44	141	43.7	949	2	US-08-539-005-50	Sequence 34, Appl
45	141	43.7	949	4	US-09-280-598-34	

ALIGNMENTS

RESULT 1
US-08-630-915A-39
; Sequence 39, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFEMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOMLAKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Misticok, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-915A-39

Alignment Scores:
Pred. No.: 5.57e-30 Length: 747
Score: 266.00 Matches: 44
Percent Similarity: 91.38% Conservative: 9
Best Local Similarity: 75.86% Mismatches: 5
Query Match: 82.35% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-630-915A-39 (1-747)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrIaGlnAsnAspAspGlnLeuAlaPhe 21
Db 553 GTATGTCAGGTGATTGCTATGTATGACTATGACCAATAATATGAAGATGAGCTCAGTTTC 612

QY 22 AsnlysglyGlnIlelleAsnValLeuAsnlysglyGlnAspProAspTyrTrpTrpIlysglyGlu 41
Db 613 TCCAAGGACAACTCATTAATGATTAAGAACAAAGATGATCTGATGTGGCCAAAGAGAG 672

QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrVallyLeuThrThrAsp 59
Db 673 ATCAACGGGGGAGCTGCTCTTCTTCTTCAAACTACGTTAAGATGACACAGAC 726

RESULT 2
US-08-630-915A-193
Sequence 193, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF SEQUENCES: 227
TITLE OF INVENTION: USING SAME
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:

LENGTH: 2873 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-630-915A-193

Alignment Scores:
Pred. No.: 3.08e-29 Length: 2873
Score: 266.00 Matches: 44
Percent Similarity: 91.38% Conservative: 9
Best Local Similarity: 75.86% Mismatches: 5
Query Match: 82.35% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-630-915A-193 (1-2873)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrIaGlnAsnAspAspGlnLeuAlaPhe 21
Db 1338 GTATGTCAGGTGATTGCTATGTATGACTATGACCAATAATATGAAGATGAGCTCAGTTTC 1397

QY 22 AsnlysglyGlnIlelleAsnValLeuAsnlysglyGlnAspProAspTyrTrpTrpIlysglyGlu 41
Db 1398 TCCAAGGACAACTCATTAATGATTAAGAACAAAGATGATCTGATGTGGCCAAAGAGAG 1457

QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrVallyLeuThrThrAsp 59
Db 1458 ATCAACGGGGGAGCTGCTCTTCTTCTTCAAACTACGTTAAGATGACACAGAC 1511

RESULT 3
US-08-475-894-3
Sequence 3, Application US/08475894
Patent No. 5641748
GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,894
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-475-894-3

Alignment Scores:
Pred. No.: 4.16e-14 Length: 1392
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12

Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 1 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-475-894-3 (1-1392)

QY 3 CysGlnValIleGlyMetYrAspYrThrAlaGlnAsnAspGluLeuAlaPheAsn 22
DB 763 TGCAGAGTATA-----TTTCCATATGAGGACAGATGATGATGACATCAAA 816
QY 23 LysGlyGlnIleIleAsnValLeuAsnLysGlu-----AspProAspTrpTrpLysGly 40
DB 817 GAAGGAGATATGATGATCTCATCAATAGAGCTGCATGACGTAAGGCTGTGTGGAAAGA 876
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyValLysLeu 56
DB 877 GAGCTGAACGGCAGACGAGCGGTTCCTCCGATTAAGTTA 924

RESULT 4

US-08-484-710-3
; Sequence 3, Application US/08484710
; Patent No. 5656438
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,710
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-484-710-3

Alignment Scores:

Pred. No.: 4.16e-14 Length: 1392
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 1 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-484-710-3 (1-1392)

QY 3 CysGlnValIleGlyMetYrAspYrThrAlaGlnAsnAspGluLeuAlaPheAsn 22
DB 763 TGCAGAGTATA-----TTTCCATATGAGGACAGATGATGATGACATCAAA 816
QY 23 LysGlyGlnIleIleAsnValLeuAsnLysGlu-----AspProAspTrpTrpLysGly 40

DB 817 GAAGGAGATATGATCTCATCAATAGAGCTGCATGACGTAAGCTGTGTGGAAAGA 876

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyValLysLeu 56
DB 877 GAGCTGAACGGCAGACGAGCGGTTCCTCCGATTAAGTTA 924

RESULT 5

US-08-484-709-3
; Sequence 3, Application US/08484709
; Patent No. 5837844
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,709
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-192
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-484-709-3

Alignment Scores:

Pred. No.: 4.16e-14 Length: 1392
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 2 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-484-709-3 (1-1392)

QY 3 CysGlnValIleGlyMetYrAspYrThrAlaGlnAsnAspGluLeuAlaPheAsn 22
DB 763 TGCAGAGTATA-----TTTCCATATGAGGACAGATGATGATGACATCAAA 816
QY 23 LysGlyGlnIleIleAsnValLeuAsnLysGlu-----AspProAspTrpTrpLysGly 40
DB 817 GAAGGAGATATGATCTCATCAATAGAGCTGCATGACGTAAGCTGTGTGGAAAGA 876
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyValLysLeu 56
DB 877 GAGCTGAACGGCAGACGAGCGGTTCCTCCGATTAAGTTA 924

RESULT 6
US-08-474-697-3
; Sequence 3, Application US/08474697
; Patent No. 6171800

```

: GENERAL INFORMATION:
: APPLICANT: Yen-Ming Hsu
: TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESS: LAHIVE & COCKFIELD
: STREET: 60 State Street, Suite 510
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474,697
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Louis Myers
: REGISTRATION NUMBER: 35,965
: REFERENCE/DOCKET NUMBER: BGP-193
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1392 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
:
: US-08-474-697-3
:
: Alignment Scores:
: Pred. No.: 4,15e-14 Length: 1392
: Score: 160.00 Matches: 29
: Percent Similarity: 73.21% Conservative: 12
: Best Local Similarity: 51.79% Mismatches: 11
: Query Match: 49.54% Indels: 4
: Gaps: 2
:
: US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-474-697-3 (1-1392)
:
: QY 3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheasn 22
:   |||:::|||||:::
: Db 763 TGCAAGTAATA-----TTTCCATATGAGGACACAGATGATGATGCAATTCACAA 816
:
: QY 23 LysGlyGlnIleIleAsnValLeuAsnLysGlu-----AappProAspTyrTrpLysGly 40
:   |||:::|||||:::|||||:::|||||:::
: Db 817 GAAGGAGATATAGTACTCTCATCAATAGAGACTGCATCGACGTRGGCTGTGGAGAGGA 876
:
: QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
:   |||:::|||||:::|||||:::|||||:::
: Db 877 GAGCTGAACGGCAGACGAGCGGCTGTCCCGATTAATTCTCGAAGTTA 924
:
: RESULT 7
: US-08-475-894-1
: Sequence 1, Application US/08475894
: Patent No. 5641748
: GENERAL INFORMATION:
: APPLICANT: Yen-Ming Hsu
: TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESS: LAHIVE & COCKFIELD
: STREET: 60 State Street, Suite 510
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/475,894
: FILING DATE:
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Louis Myers
: REGISTRATION NUMBER: 35,965
: REFERENCE/DOCKET NUMBER: BGP-191
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1659 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
:
: US-08-475-894-1
:
: Alignment Scores:
: Pred. No.: 5,19e-14 Length: 1659
: Score: 160.00 Matches: 29
: Percent Similarity: 73.21% Conservative: 12
: Best Local Similarity: 51.79% Mismatches: 11
: Query Match: 49.54% Indels: 4
: Gaps: 2
:
: US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-475-894-1 (1-1659)
:
: QY 3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheasn 22
:   |||:::|||||:::
: Db 1030 TGCAAGTAATA-----TTTCCATATGAGGACACAGATGATGATGCAATTCACAA 1083
:
: QY 23 LysGlyGlnIleIleAsnValLeuAsnLysGlu-----AappProAspTyrTrpLysGly 40
:   |||:::|||||:::|||||:::|||||:::
: Db 1084 GAAGGAGATATAGTACTCTCATCAATAGAGACTGCATCGACGTRGGCTGTGGAGAGGA 1143
:
: QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
:   |||:::|||||:::|||||:::|||||:::
: Db 1144 GAGCTGAACGGCAGACGAGCGGCTGTCCCGATTAATTCTCGAAGTTA 1191
:
: RESULT 8
: US-08-484-710-1
: Sequence 1, Application US/08484710
: Patent No. 5656438
: GENERAL INFORMATION:
: APPLICANT: Yen-Ming Hsu
: TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESS: LAHIVE & COCKFIELD
: STREET: 60 State Street, Suite 510
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,710
: FILING DATE:
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Louis Myers

```

```

; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-484-710-1

Alignment Scores:
Pred. No.: 5.19e-14 Length: 1659
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 1 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-484-710-1 (1-1659)

QY 3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheAsn 22
Db 1030 TCGAAGTATA-----TTCCATATGAGGCACAGAAATGATGATTAATTCACATCAAA 1083

QY 23 LysGlyGlnIleLeuAsnValLeuAsnLysGlu-----AspProAspTyrPrlPylsGly 40
Db 1084 GAAGGAGATATAGTCTCATCATCAATAGAGACTGCATCGACGTGAGCTGGGAAAGA 1143

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
Db 1144 GAGCTGAACGCGACAGACGCGCTTCCCGATTAACCTGCTGAAGTTA 1191

RESULT 9
US-08-484-709-1
; Sequence 1, Application US/08484709
; Patent No. 5837844
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,709
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-192
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
```

```

; MOLECULE TYPE: cDNA
; US-08-484-709-1

Alignment Scores:
Pred. No.: 5.19e-14 Length: 1659
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 2 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-484-709-1 (1-1659)

QY 3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheAsn 22
Db 1030 TCGAAGTATA-----TTCCATATGAGGCACAGAAATGATGATTAATTCACATCAAA 1083

QY 23 LysGlyGlnIleLeuAsnValLeuAsnLysGlu-----AspProAspTyrPrlPylsGly 40
Db 1084 GAAGGAGATATAGTCTCATCATCAATAGAGACTGCATCGACGTGAGCTGGGAAAGA 1143

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
Db 1144 GAGCTGAACGCGACAGACGCGCTTCCCGATTAACCTGCTGAAGTTA 1191

RESULT 10
US-08-474-697-1
; Sequence 1, Application US/08474697
; Patent No. 6171800
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,697
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-474-697-1

Alignment Scores:
Pred. No.: 5.19e-14 Length: 1659
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 3 Gaps: 2
```

	US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-474-697-1 (1-1659)
OY	3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAppGluLeuAlaPheAsn 22 ::: :::
Dd	1030 TGCAGATTAATA-----TTTCCTATACAGCGCACAGATCATGTGAATTGACAAATCMAA 1083
OY	23 LysGlyGlnIleIleIleAsnValIleAsnLysGlu-----AspProAspTrpTrpLysGly 40 ::: :::
Dd	1084 GAAGGAGATNATAGACTCTCAATCAATTAAGGACTGCATCGACGTGAGCTGGTGGAAGA 1143
OY	41 GluAlaIsngIysGlnValGlyLeuPheProSerAsnTyrValLysLeu 56 ::: :::
Dd	1144 GAGCTGAACGCCAGACGAGCGCTTCCCGCATACTTCGTGAAGTTA 1191
RESULT 11	
US-08-671-354-1	
/ Sequence 1, Application US/08671354	
/ Patent No. 6423824	
GENERAL INFORMATION:	
APPLICANT: Hsu, Yen-Ming	
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY	
NUMBER OF SEQUENCES: 2	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: Fish & Richardson P.C.	
STREET: 225 Franklin Street	
CITY: Boston	
STATE: MA	
COUNTRY: USA	
ZIP: 02110-2804	
COMPUTER READABLE FORM:	
MEDIUM TYPE: Discrete	
COMPUTER: IBM Compatible	
OPERATING SYSTEM: Windows 95	
SOFTWARE: FastSeq for Windows Version 2.0b	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: US/08/671.354	
FILING DATE: 27-JUN-1996	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: US 08/486,344	
FILING DATE: 07-JUN-1995	
APPLICATION NUMBER: US 08/484,709	
FILING DATE: 07-JUN-1995	
APPLICATION NUMBER: US 08/475,894	
FILING DATE: 07-JAN-1995	
APPLICATION NUMBER: US 08/475,710	
FILING DATE: 07-JAN-1995	
APPLICATION NUMBER: US 08/474,697	
FILING DATE: 07-JUN-1995	
ATTORNEY/AGENT INFORMATION:	
NAME: Myers, Louis	
REGISTRATION NUMBER: 35,965	
REFERENCE/DOCKET NUMBER: 10274/009005	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: 617/542-5070	
TELEFAX: 617/542-8906	
TELEX: 200154	
INFORMATION FOR SEQ ID NO: 1:	
SEQUENCE CHARACTERISTICS:	
LENGTH: 3143 base pairs	
TYPE: nucleic acid	
STRANDEDNESS: single	
TOPOLOGY: linear	
MOLECULE TYPE: cDNA	
FEATURE:	
NAME/KEY: Coding Sequence	
LOCATION: 436...2412	
US-08-671-354-1	
Alignment Scores:	
Pred. No.: 1.17e-13	Length: 3143
Score: 160.00	Matches: 29
Percent Similarity: 73.21%	Conservative: 12
Best Local Similarity: 51.79%	Mismatches: 11

```

Query Match: 49.54% Indels: 4
DB: 4 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-671-354-1 (1-3143)
QY      3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheAsn 22
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       1231 TGCAAAGATTA-----TTTCCTATTACGAGCAGCAATGATGATGACAATTCANA 128
QY      23 LysGluGlnIleIleAsnValLeuAsnLysGlu-----AspProAspTrpTrpLysGly 40
        ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       1285 GAAGAGCATATAGCACACTCTCATCAATATGAAGCATGCACTGATGAGCGTGCGGAAGA 134
QY      41 GluValAsnGlyGlnValAlaGlyLeuPheProSerAsnTyrValLeu 56
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       1345 GAGCTGAACGGCAACAGCAGGCGGTGCCCCGATTAACCTTGTAAGATT 1352

RESULT 12
US-08-475-894-5
Sequence 5, Application US/08475894
Patent No. 5641748
GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,894
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGE-191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO.: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1045 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-475-894-5

Alignment Scores:
Pred. No.: 1,11e-13 Length: 1045
Score: 156.00 Matches: 26
Percent Similarity: 73.08% Conservative: 12
Best Local Similarity: 50.00% Mismatches: 14
Query Match: 48.30% Indels: 0
Gaps: 0
DB: 1

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-475-894-5 (1-1045)
QY      4 GlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheAsnLys 23
        ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       118 GAGGCAATATGAGAGTTGACTACACAGGCCACGACGATGATGAGTGCAGATCAGCGTG 177
QY      24 GlyGlnIleIleAsnValLeuAsnLysGluAspProAspTrpTrpLysGlyGlnValAsn 43
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```


MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,697
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1045 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-474-697-5

Alignment Scores:
Pred. No.: 1,11e-13 Length: 1045
Score: 156.00 Matches: 26
Percent Similarity: 73.08% Conservative: 12
Best Local Similarity: 50.00% Mismatches: 14
Query Match: 48.30% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-474-697-5 (1-1045)

QY 4 GlnVal11eg1yMeTyrrAsPTyrrThra1aGlnasnAspAspGluLeuAlaPheAsnLys 23
Db 118 GAGGCCATAGTGAGATTGACTACCCAGGCCCGACGATGATGAGCTGACGATCAGCGTG 177
QY 24 G1yGln1n1e1leAsnValLeuAsnLysGluAspProAspTrrP1ySg1yGluValasn 43
Db 178 GGTGAATTCATCCACCAACATCAGAGAGGATGAGGCTGTGTGGAGGAGACAGATCAAC 237
QY 44 G1yGlnValG1yLeuPheProSerAsnTyrrValys 55
Db 238 GGCAGAGAGGTTGTTCCTGACACTTGTAGA 273

Search completed: July 1, 2004, 19:54:00
Job time : 45.3138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 181.007 Seconds

(without alignments)

1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_1080_1138
Perfect score: 323
Sequence: 1 AVOGVIGMDYTRQNDELAA.....GEVNGVGLPRMYVLTLD 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODER=frame+ p2n.model -DEV=xlp
-Q/cg2_1/USPTO.spool.p/US09720934/runtac_30062004_064541_13481/app_query.fasta_1.1386
-DB=Published Applications NA -QFMT=fastap -SUFFIX=trmb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=Pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODER=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09720934 @CGN_1_1500 @runtac_30062004_064541_13481
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOP -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database: Published Applications NA:*

1: /cg2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:*
3: /cg2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:*
10: /cg2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
11: /cg2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:*
12: /cg2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
13: /cg2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
14: /cg2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
15: /cg2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
16: /cg2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:*
17: /cg2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
18: /cg2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
19: /cg2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	323	100.0	2067	16	US-10-264-049-887	Sequence 887, App
2	266	82.4	747	9	US-09-879-957-39	Sequence 39, Appl
3	266	82.4	2873	9	US-09-879-957-193	Sequence 193, App
4	266	82.4	3746	11	US-09-764-875-176	Sequence 176, App
5	266	82.4	4053	13	US-10-342-887-1882	Sequence 1882, Ap
6	266	82.4	4053	13	US-10-172-118-1882	Sequence 1882, Ap
7	266	82.4	4210	9	US-09-764-868-125	Sequence 125, Ap
8	266	82.4	5828	13	US-10-398-885A-15	Sequence 15, Appl
9	160	49.5	386	16	US-10-085-783A-37196	Sequence 37196, A
10	160	49.5	400	13	US-10-242-535A-22639	Sequence 22639, A
11	160	49.5	400	13	US-10-085-783A-22639	Sequence 22639, A
12	160	49.5	400	16	US-10-242-535A-22639	Sequence 22639, A
13	160	49.5	1215	17	US-10-648-593-61	Sequence 61, Appl
14	160	49.5	3143	14	US-10-144-621-1	Sequence 1, Appl1
15	160	49.5	3348	13	US-10-342-887-187	Sequence 187, App
16	160	49.5	3348	13	US-10-172-118-187	Sequence 187, App
17	159.5	49.4	2055	16	US-10-085-117-356	Sequence 356, App
18	159.5	49.4	4341	16	US-10-085-117-356	Sequence 359, App
19	158.5	49.1	7696	16	US-10-264-237-2837	Sequence 2837, App
20	158.5	49.1	325348	16	US-10-085-117-358	Sequence 358, App
21	158	48.9	184	13	US-10-085-783A-26214	Sequence 26214, A
22	158	48.9	184	16	US-10-242-535A-26214	Sequence 26214, A
23	158	48.9	2967	15	US-10-207-655-197	Sequence 197, App
24	155	48.0	501	9	US-09-764-868-122	Sequence 122, App
25	155	48.0	501	11	US-09-764-875-136	Sequence 136, App
26	155	48.0	501	11	US-09-764-875-444	Sequence 444, App
27	154.5	47.8	2070	16	US-10-085-117-357	Sequence 357, App
28	154.5	47.8	2841	16	US-10-085-117-356	Sequence 356, App
29	154	47.7	1803	9	US-09-962-832-6	Sequence 6, Appl1
30	154	47.7	1803	15	US-10-171-8581-43	Sequence 43, Appl1
31	153.5	47.5	194945	16	US-10-085-117-355	Sequence 355, App
32	152	47.1	548	10	US-09-918-995-28621	Sequence 28621, A
33	152	47.1	813	17	US-10-641-643-1238	Sequence 1238, Ap
34	152	47.1	984	10	US-09-765-998A-5	Sequence 5, Appl1
35	152	47.1	1109	10	US-09-960-706-713	Sequence 713, App
36	152	47.1	1109	15	US-10-007-926A-232	Sequence 232, App
37	152	47.1	1109	15	US-10-327-509-1	Sequence 1, Appl1
38	150	46.4	1661	15	US-10-187-148-2	Sequence 2, Appl1
39	148	45.8	1710	9	US-09-879-957-189	Sequence 189, App
40	146	45.2	476	10	US-09-918-995-35422	Sequence 35422, A
41	146	45.2	500	13	US-10-085-783A-17169	Sequence 17169, A
42	146	45.2	500	16	US-10-242-535A-17169	Sequence 17169, A
43	146	45.2	1889	9	US-09-925-301-274	Sequence 274, App
44	145	44.9	182	15	US-10-029-386-26602	Sequence 26602, A
45	145	44.9	550	15	US-10-029-386-12902	Sequence 12902, A

ALIGNMENTS

RESULT 1
US-10-264-049-887
Sequence 887, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Biase et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PAL33PI
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 887
LENGTH: 2067
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (5)..(5)

OTHER INFORMATION: n equals a,t,b, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2058) .. (2058)
OTHER INFORMATION: n equals a,t,b, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2063) .. (2063)
OTHER INFORMATION: n equals a,t,b, or c
US-10-264-049-887

Alignment Scores:

Pred. No.: 4,1e-41 Length: 2067
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-264-049-887 (1-2067)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
DB 299 GCAAGTGTCCAGGATGATGGAGTGAACGACTACACCGGACAAATGACGATGAGCTGGCC 358
QY 21 PheAsnIysGlyGlnIleIleAsnValLeuAsnIysGluAspProAspTyrTrpIysGly 40
DB 359 TTCAACAAAGGCGCATCATCAACCTCTCAACAAAGGAGACCCCTGACTGGTGAAAGGA 418
QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValIysLeuThrThrasp 59
DB 419 GAAGTCATGACGACAGTGGGCTCTCCCATCCATTATGTGAAGCTACACAGAC 475

RESULT 2

US-09-879-957-39
Sequence 39, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
HOFFMAN, No. US20020034755A1h
KAY, Brian K.
FOWLES, Dana M.
MCCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESS: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/879,957

FILING DATE: 13-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/630,915

FILING DATE: 03-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-879-957-39

Alignment Scores:

Pred. No.: 1.37e-32 Length: 747
Score: 266.00 Matches: 44
Percent Similarity: 91.38% Conservative: 5
Best Local Similarity: 75.86% Mismatches: 0
Query Match: 82.35% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-09-879-957-39 (1-747)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21
DB 553 GTATGTCAGTGAATGCTATGTATGACATACAGCAATATGAAGATGAGCTCAGTTTC 612
QY 22 AsnIysGlyGlnIleIleAsnValLeuAsnIysGluAspProAspTyrTrpIysGlyGlu 41
DB 613 TCCAAGGACACATCATTAATGTTATGAACAAAGTATCTCTGATGGTGGAAGAG 672
QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValIysLeuThrThrasp 59
DB 673 ATCAACGGCGTACGCTGCTCTTCTTCTTCAACACTACGTTAAGATGACACAGAC 726

RESULT 3

US-09-879-957-193
Sequence 193, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
HOFFMAN, No. US20020034755A1h
KAY, Brian K.
FOWLES, Dana M.
MCCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESS: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/879,957

FILING DATE: 13-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/630,915

FILING DATE: 03-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

```

;
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2873 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 193:
US-09-879-957-193

Alignment Scores:
Pred. No.:      8,45e-32      Length:      2873
Score:          266.00      Matches:      44
Percent Similarity: 91.38%      Conservative: 9
Best Local Similarity: 75.86%      Mismatches: 5
Query Match:      82.35%      Indels:      0
DB:               9          Gaps:      0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-09-879-957-193 (1-2873)

QY      2 ValCyGgInVal1lleglMetYrAspTYrThAlaGlnAsnAspAspGluLeuAlaPhe 21
      1338 GATATGTCAGGTGATTGCTATGATGACTATGCAGCAATATATGAAATGATGAGCTCAGTTTC 1397
      22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrPlysglyGlu 41
      1398 TCCAAGGAGCAACTCTATTATGTTATGACCAAGATATCTGATTTGGTGGCAAGAGAG 1457
      42 ValAsnGlyGlnValGlyLeuPheProSerAsnTYrValIlysLeuThrThraap 59
      1458 ATCAACGGGGTGACTGCTCTTCTTCTTCAAACTAGCTTAAGATGACGACAGAC 1511
      RESULT 4
      US-09-764-875-176
      ; Sequence 176, Application US/09764875
      ; Publication No. US20040018969A1
      ; GENERAL INFORMATION:
      ; APPLICANT: Rosen et al.
      ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
      ; FILE REFERENCE: PJ202
      ; CURRENT APPLICATION NUMBER: US/09/764,875
      ; CURRENT FILING DATE: 2001-01-17
      ; Prior application data removed - consult PALM or file wrapper
      ; NUMBER OF SEQ ID NOS: 1249
      ; SOFTWARE: PatentIn Ver. 2.0
      ; SEQ ID NO 176
      ; LENGTH: 3746
      ; TYPE: DNA
      ; ORGANISM: Homo sapiens
      US-09-764-875-176

Alignment Scores:
Pred. No.:      1.21e-31      Length:      3746
Score:          266.00      Matches:      44
Percent Similarity: 91.38%      Conservative: 9
Best Local Similarity: 75.86%      Mismatches: 5
Query Match:      82.35%      Indels:      0
DB:               11         Gaps:      0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-09-764-875-176 (1-3746)

QY      2 ValCyGgInVal1lleglMetYrAspTYrThAlaGlnAsnAspAspGluLeuAlaPhe 21
      1264 GATATGTCAGGTGATTGCTATGATGACTATGCAGCAATATATGAAATGATGAGCTCAGTTTC 1323
      22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrPlysglyGlu 41
      1324 TCCAAGGAGCAACTCTATTATGTTATGACCAAGATATCTGATTTGGTGGCAAGAGAG 1383
      42 ValAsnGlyGlnValGlyLeuPheProSerAsnTYrValIlysLeuThrThraap 59
      1384 ATCAACGGGGTGACTGCTCTTCTTCTTCAAACTAGCTTAAGATGACGACAGAC 1437
```

```

      RESULT 5
      US-10-342-887-1882
      ; Sequence 1882, Application US/10342887
      ; Publication No. US20040058340A1
      ; GENERAL INFORMATION:
      ; APPLICANT: Dai, Hongyue
      ; APPLICANT: He, Yudong
      ; APPLICANT: Linsley, Peter S.
      ; APPLICANT: Mao, Mao
      ; APPLICANT: Roberts, Christopher J.
      ; APPLICANT: Van 't Veer, Laura Johanna
      ; APPLICANT: Van de Vijver, Marc J.
      ; APPLICANT: Bernards, Rene
      ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
      ; FILE REFERENCE: 9301-188-999
      ; CURRENT APPLICATION NUMBER: US/10/342,887
      ; CURRENT FILING DATE: 2003-01-15
      ; PRIOR APPLICATION NUMBER: 60/298,918
      ; PRIOR FILING DATE: 2001-06-18
      ; PRIOR APPLICATION NUMBER: 60/380,710
      ; PRIOR FILING DATE: 2002-05-14
      ; PRIOR APPLICATION NUMBER: 10/172,118
      ; PRIOR FILING DATE: 2002-06-14
      ; NUMBER OF SEQ ID NOS: 2639
      ; SEQ ID NO 1882
      ; LENGTH: 4053
      ; TYPE: DNA
      ; ORGANISM: Homo sapiens
      US-10-342-887-1882

Alignment Scores:
Pred. No.:      1.35e-31      Length:      4053
Score:          266.00      Matches:      44
Percent Similarity: 91.38%      Conservative: 9
Best Local Similarity: 75.86%      Mismatches: 5
Query Match:      82.35%      Indels:      0
DB:               13         Gaps:      0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-342-887-1882 (1-4053)

QY      2 ValCyGgInVal1lleglMetYrAspTYrThAlaGlnAsnAspAspGluLeuAlaPhe 21
      553 GATATGTCAGGTGATTGCTATGATGACTATGCAGCAATATATGAAATGATGAGCTCAGTTTC 612
      22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrPlysglyGlu 41
      613 TCCAAGGAGCAACTCTATTATGTTATGACCAAGATATCTGATTTGGTGGCAAGAGAG 672
      42 ValAsnGlyGlnValGlyLeuPheProSerAsnTYrValIlysLeuThrThraap 59
      673 ATCAACGGGGTGACTGCTCTTCTTCTTCAAACTAGCTTAAGATGACGACAGAC 726
      RESULT 6
      US-10-172-118-1882
      ; Sequence 1882, Application US/10172118
      ; Publication No. US20030224374A1
      ; GENERAL INFORMATION:
      ; APPLICANT: Dai, Hongyue
      ; APPLICANT: He, Yudong
      ; APPLICANT: Linsley, Peter
      ; APPLICANT: Mao, Mao
      ; APPLICANT: Roberts, Chris
      ; APPLICANT: Van 't Veer, Laura
      ; APPLICANT: Van de Vijver, Marc
      ; APPLICANT: Bernards, Rene
      ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
      ; FILE REFERENCE: 9301-175-999
      ; CURRENT APPLICATION NUMBER: US/10/172,118
      ; CURRENT FILING DATE: 2002-06-14
      ; PRIOR APPLICATION NUMBER: 60/380,770
      ; PRIOR FILING DATE: 2002-05-14
      ; NUMBER OF SEQ ID NOS: 2699
```

```
; SEQ ID NO 1882
; LENGTH: 4053
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: U61167
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1882

Alignment Scores:
Pred. No.: 1,35e-31 Length: 4053
Score: 266.00 Matches: 44
Percent Similarity: 91.38% Conservative: 9
Best Local Similarity: 75.86% Mismatches: 5
Query Match: 82.35% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-172-118-1882 (1-4053)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21
DB 553 GTATGTCAGGAGTTCCTATGATGACTATGACGCAAAATATGAGATGAGCTCAGTTTC 612
QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGlu 41
DB 613 TCCAAGGACAACTCATTAATGATATGAAACAAGATGATCCTGATGGTGGCAAGAGAG 672
QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThraSp 59
DB 673 ATCAACGGGGTGACTGCTCTCTCTTCTTCAAACTAGTTAAGATGACGACAGAC 726

RESULT 7
US-09-764-868-125
; Sequence 125, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT732
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - refer to PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 125
; LENGTH: 4210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-125

Alignment Scores:
Pred. No.: 1.42e-31 Length: 4210
Score: 266.00 Matches: 44
Percent Similarity: 91.38% Conservative: 9
Best Local Similarity: 75.86% Mismatches: 5
Query Match: 82.35% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-09-764-868-125 (1-4210)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21
DB 1715 GTATGTCAGGAGTTCCTATGATGACTATGACGCAAAATATGAGATGAGCTCAGTTTC 1774
QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGlu 41
DB 1775 TCCAAGGACAACTCATTAATGATATGAAACAAGATGATCCTGATGGTGGCAAGAGAG 1834
QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThraSp 59
DB 1835 ATCAACGGGGTGACTGCTCTCTCTTCTTCAAACTAGTTAAGATGACGACAGAC 1888

RESULT 8
```

```
US-10-398-885A-15
; Sequence 15, Application US/10398885A
; Publication No. US20040053282A1
; GENERAL INFORMATION:
; APPLICANT: Sugita, Yuji
; APPLICANT: Hashida, Ryochi
; APPLICANT: Ogawa, Kaoru
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Obayashi, Masaya
; APPLICANT: Saito, Hirohisa
; APPLICANT: Takahashi, Eiki
; TITLE OF INVENTION: Method of Testing For Allergic Diseases
; FILE REFERENCE: SHIMIZU-07907
; CURRENT APPLICATION NUMBER: US/10/398,885A
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: PCT/JP01/08937
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: JP 2000-314093
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 5828
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: CDS
; LOCATION: (7)..(5052)
; OTHER INFORMATION:
US-10-398-885A-15

Alignment Scores:
Pred. No.: 2.2e-31 Length: 5828
Score: 266.00 Matches: 44
Percent Similarity: 91.38% Conservative: 9
Best Local Similarity: 75.86% Mismatches: 5
Query Match: 82.35% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-398-885A-15 (1-5828)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21
DB 3343 GTATGTCAGGAGTTCCTATGATGACTATGACGCAAAATATGAGATGAGCTCAGTTTC 3402
QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGlu 41
DB 3403 TCCAAGGACAACTCATTAATGATATGAAACAAGATGATCCTGATGGTGGCAAGAGAG 3462
QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThraSp 59
DB 3463 ATCAACGGGGTGACTGCTCTCTCTTCTTCAAACTAGTTAAGATGACGACAGAC 3516

RESULT 9
US-10-085-783A-37196
; Sequence 37196, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
```

Qy	3	CysGlnValIleIleGlyMetCysIleAspTyrThrAlaGlnAsnAspAspGluIleValAlaPheAsn	22
Db	137	TGCAAGTATATA-----TTTCCATATAGACGACAGATGATGATGATGACATCAAA	190
Qy	23	LysGlyGlnIleIleLeuValLeuAsnLysGlu-----AspProAspTyrTyrLysGly	40
Db	191	GAGGAGATATATAGTACTCTCTCATATAGAGACTGCATCGACGTAAGGCTGGTGGGAAAGA	25
Qy	41	GluValAsnGlyGlnValGlyIleuPheProSerAsnTyrValLysLeu	56
Db	251	GAGCTGAACGGCAGACGAGCGCTGTCCCGCATTACTTCGTAAGTTA	298
RESULT 11			
US-10-085-783A-22639			
Sequence 22639, Application US/10085783A			
Publication No. US20040037841A1			
GENERAL INFORMATION:			
APPLICANT: ChondroGene Inc.			
APPLICANT: Ilex, C.C.			
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis			
FILE REFERENCE: 4231/2002			
CURRENT APPLICATION NUMBER: US/10/085,783A			
CURRENT FILING DATE: 2002-02-28			
PRIOR APPLICATION NUMBER: US 60/305,340			
PRIOR FILING DATE: 2001-07-13			
PRIOR APPLICATION NUMBER: US 60/275,017			
PRIOR FILING DATE: 2001-03-12			
PRIOR APPLICATION NUMBER: US 60/271,955			
PRIOR FILING DATE: 2001-02-28			
NUMBER OF SEQ. ID NOS: 58994			
SOFTWARE: PatentIn version 3.2			
SEQ ID NO 22639			
LENGTH: 400			
TYPE: DNA			
ORGANISM: Human			
FEATURE:			
NAME/KEY: misc_feature			
LOCATION: (9)..(9)			
OTHER INFORMATION: n is a, c, g, or t			
FEATURE:			
NAME/KEY: misc_feature			
LOCATION: (16)..(16)			
OTHER INFORMATION: n is a, c, g, or t			
FEATURE:			
NAME/KEY: misc_feature			
LOCATION: (68)..(68)			
OTHER INFORMATION: n is a, c, g, or t			
FEATURE:			
NAME/KEY: misc_feature			
LOCATION: (75)..(75)			
OTHER INFORMATION: n is a, c, g, or t			
US-10-085-783A-22639			
Alignment Scores:			
Pred. No.: 5.43e-16			
Score: 160.00			
Percent Similarity: 73.21%			
Best Local Similarity: 51.79%			
Query Match: 49.54%			
DB: 13			
Gaps: 2			
US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-085-783A-22639 (1-400)			
Qy	3	CysGlnValIleIleGlyMetCysIleAspTyrThrAlaGlnAsnAspAspGluIleValAlaPheAsn	22
Db	159	TGCAAGTATATA-----TTTCCATATAGAGCAGACAGATGATGATGATGACATCAAA	212
Qy	23	LysGlyGlnIleIleLeuValLeuAsnLysGlu-----AspProAspTyrTyrLysGly	40
Db	213	GAGGAGATATATAGTACTCTCTCATATAGAGACTGCATCGACGTAAGGCTGGTGGGAAAGA	272
Qy	41	GluValAsnGlyGlnValGlyIleuPheProSerAsnTyrValLysLeu	56

```
Db 273 GAGCTGAACGGACAGAGCGCGTGTCCCGATTAACCTTGGAAGTTA 320
RESULT 12
US-10-242-535A-22639
; Sequence 22639, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22639
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (68)..(68)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (75)..(75)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-22639
Alignment Scores:
Pred. No.: 5.43e-16 Length: 400
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 16 Gaps: 2
US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-242-535A-22639 (1-400)
QY 3 CysGlnValIleIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaIleAsn 22
Db 159 TGCAGAGTAATA-----TTTCCATATGAGGACAGCAAGATGATGATGAATTCACATCAA 212
QY 23 LysGlyGlnIleIleAsnValLeuAsnLysGlu-----AspProAspTyrTrpLysGly 40
Db 213 GAAGGAGATATAGTACCTCTCATCAATAAGACCTGCATCGACGTAGGCTGGTGGGAGAGA 272
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
Db 273 GAGCTGAACGGACAGAGCGCGTGTCCCGATTAACCTTGGAAGTTA 320
RESULT 13
US-10-648-593-61
; Sequence 61, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-61
Alignment Scores:
Pred. No.: 2.43e-15 Length: 1215
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 17 Gaps: 2
US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-648-593-61 (1-1215)
QY 3 CysGlnValIleIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaIleAsn 22
Db 31 TGCAGAGTAATA-----TTTCCATATGAGGACAGCAAGATGATGATGAATTCACATCAA 84
QY 23 LysGlyGlnIleIleAsnValLeuAsnLysGlu-----AspProAspTyrTrpLysGly 40
Db 85 GAAGGAGATATAGTACCTCTCATCAATAAGACCTGCATCGACGTAGGCTGGTGGGAGAGA 144
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
Db 145 GAGCTGAACGGACAGAGCGCGTGTCCCGATTAACCTTGGAAGTTA 192
RESULT 14
US-10-144-621-1
; Sequence 1, Application US/10144621
; Publication No. US20020172671A1
; GENERAL INFORMATION:
; APPLICANT: Heu, Yen-Ming
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/144,621
FILING DATE: 13-May-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/671,354
FILING DATE: 27-JUN-1996
APPLICATION NUMBER: US 08/486,344
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/484,709
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/475,894
FILING DATE: 07-JAN-1995
APPLICATION NUMBER: US 08/475,710
FILING DATE: 07-JAN-1995
```

```

; APPLICATION NUMBER: US 08/474,697
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10274/009005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3143 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 436..2412
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-144-621-1
;

Alignment Scores:
Pred. No.: 8.77e-15 Length: 3143
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 14 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-144-621-1 (1-3143)
QY 3 CysGlnValIleGlyMetYrAspIYrThrAlaGlnAsnAspAspGlnLeuAlaPheAsn 22
|||:::|||||:::
Db 1231 TGCAGAGTAATA-----TTTCCATATGAGGCACAGAAATGATGATTAATGACAAATCAAA 1284
|||:::|||||:::

QY 23 LysGlyGlnIleIleAsnValLeuAsnLysGlu-----AspProAspTyrPrtPlySgly 40
|||:::|||||:::
Db 1285 GAAAGAGATATTAATCTCTCATCAATTAAGAGACTGCATCGACGTAGCGTGGTGGAAAGA 1344
|||:::|||||:::

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValIysLeu 56
|||:::|||||:::
Db 1345 GAGCTGAACGGCAGACGAGCGCGTGTCCCGAATACTTGTGTAAGTTA 1392
|||:::|||||:::

RESULT 15
US-10-342-887-187
; Sequence 187, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 187
; LENGTH: 3348
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```

US-10-342-887-187

Alignment Scores:
Pred. No.: 9.55e-15 Length: 3348
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 13 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-342-887-187 (1-3348)
QY 3 CysGlnValIleGlyMetYrAspIYrThrAlaGlnAsnAspAspGlnLeuAlaPheAsn 22
|||:::|||||:::
Db 1105 TGCAGAGTAATA-----TTTCCATATGAGGCACAGAAATGATGATTAATGACAAATCAAA 1158
|||:::|||||:::

QY 23 LysGlyGlnIleIleAsnValLeuAsnLysGlu-----AspProAspTyrPrtPlySgly 40
|||:::|||||:::
Db 1159 GAAAGAGATATTAATCTCTCATCAATTAAGAGACTGCATCGACGTAGCGTGGTGGAAAGA 1218
|||:::|||||:::

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValIysLeu 56
|||:::|||||:::
Db 1219 GAGCTGAACGGCAGACGAGCGCGTGTCCCGAATACTTGTGTAAGTTA 1266
|||:::|||||:::

Search completed: July 1, 2004, 20:16:15
Job time : 188.007 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 1774.84 Seconds

(Without alignments) 992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_1080_1138
Perfect score: 323
Sequence: 1 AVCGVIGMVDYTAQNDELAA.....GEVNGGVGLPPSNVYKLTDD 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgmt2.1/USPTO.spool_P/US09720934/runat_30062004_064540_13442/app_query.fasta_1.1386
-DB=EST -QFMT=fastp -SUFFIX=rst -MIMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-DOCAIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEADSIZE=500 -MINLEN=2000000000
-USER=US09720934.@CCN_1_1.12431.@runat_30062004_064540_13442 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -IONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estlm.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_lnv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rdg.*
26: em_gss_prg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	323	100.0	487	9	AI151142
2	323	100.0	545	14	CB161912
3	323	100.0	575	14	CB163763
4	323	100.0	583	14	CB156803
5	323	100.0	591	14	CB215044
6	323	100.0	1062	13	BO898718
7	323	100.0	1064	10	BE888832
8	319	98.8	363	14	H66900
9	317	98.1	581	12	BG894858
10	317	98.1	612	12	BG894850
11	315	97.5	693	13	BU321701
12	313	96.9	623	13	BU056026
13	313	96.9	634	13	BU054334
14	313	96.9	782	12	BG922829
15	313	96.9	782	13	BU614534
16	313	96.9	906	13	BU519029
17	313	96.9	2589	11	AK082606
18	313	96.9	5385	11	BC062938
19	312	96.6	538	14	CA390605
20	312	96.6	1073	13	BQ432912
21	308	95.4	406	9	AV590688
22	308	95.4	594	14	CB172533
23	308	95.4	612	14	CB435542
24	308	95.4	629	14	CF768246
25	306	94.7	603	9	AI594919
26	301.5	93.3	521	12	BG876696
27	295	91.3	591	12	BU639774
28	295	91.3	713	14	CB209507
29	295	91.3	898	14	CA987804
30	293	90.7	749	13	BU963831
31	293	90.7	830	14	CB988722
32	283	87.6	674	13	BU054591
33	281	87.0	539	9	AL915730
34	276	85.4	1036	29	CNS02AR3
35	272	84.2	552	12	BM247316
36	268	83.0	684	14	CA365705
37	266	82.4	225	14	CF135884
38	266	82.4	372	10	BP918427
39	266	82.4	455	12	BM757984
40	266	82.4	458	12	BI031994
41	266	82.4	517	12	BM750308
42	266	82.4	640	12	BM758638
43	265	82.4	702	13	BX502192
44	265	82.0	549	12	BM233591
45	265	82.0	576	10	AW631635

ALIGNMENTS

RESULT 1
AI151142/c
LOCUS AI151142
DEFINITION qc87e04.x1 Soares pregnant uterus Nbhpu Homo sapiens CDNA clone
IMAGE:171214 3' similar to SW:SH17 HUMAN Q15811 SH3
DOMAIN-CONTAINING PROTEIN SH3p17. ; mRNA sequence.
ACCESSION AI151142
VERSION AI151142.1 GI:3679611
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 487)
 AUTHORS NC1-GAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 518 Std Error: 0.00
 Seq primer: -40m13 fwd. fr from AmerSham
 High quality sequence stop: 444.
 Location/Qualifiers
 source
 1..487
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1721214"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pT73-Pac; Site 1: Not I;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5',
 AACTGGAAGATTGCGCGCCCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by M. Fatima Bonaldo."

ORIGIN
 Alignment Scores:
 Pred. No.: 2.83e-36 Length: 487
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x A1151142 (1-487)

QY 1 A1A1A1CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
 Db 430 GCAGTGTCCAGGAGTGGATGGATGACACTACACCGCGCAGAAATGACGATGAGCTGGCC 371

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40
 Db 370 TTCAACAAGGGCCAGATCATCACTCTCTCAACAAGGAGGACCCTGACTGGTGGAAAGA 311

QY 41 GluValaAnglyGlnValaGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
 Db 310 GAAGTCATGACAGATGGGGCTCTTCCCATCCCAATTATGTGAAGCTGACCAACAGAC 254

RESULT 2
 LOCUS CB161912 545 bp mRNA linear EST 30-JAN-2003
 DEFINITION K-EST0221969 L17N670205n1 Homo sapiens cDNA clone
 L17N670205n1-41-H01 5', mRNA sequence.
 ACCESSION CB161912
 VERSION CB161912.1 GI:28148038
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 545)
 AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.U., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 41 row: H column: 01
 High quality sequence stop: 545.
 Location/Qualifiers
 source
 1..545
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L17N670205n1-41-H01"
 /sex="F"
 /lab_host="TOP10F"
 /clone_id="L17N670205n1"
 /note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI;
 Site 2: NotI; The library was contributed by the Soares
 laboratory and it was constructed as described by Bonaldo,
 M.F., Lemon, G. and Soares, M.B. (1996), Genome Research
 6(9): 791-806. RNA was prepared from harvested cell
 culture."

ORIGIN
 Alignment Scores:
 Pred. No.: 3.34e-36 Length: 545
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x CB161912 (1-545)

QY 1 A1A1A1CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
 Db 361 GCAGTGTCCAGGAGTGGATGGATGACACTACACCGCGCAGAAATGACGATGAGCTGGCC 420

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40
 Db 421 TTCAACAAGGGCCAGATCATCACTCTCTCAACAAGGAGGACCCTGACTGGTGGAAAGA 480

QY 41 GluValaAnglyGlnValaGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
 Db 481 GAAGTCATGACAGATGGGGCTCTTCCCATCCCAATTATGTGAAGCTGACCAACAGAC 537

RESULT 3
 LOCUS CB163763 575 bp mRNA linear EST 30-JAN-2003
 DEFINITION K-EST0224708 L17N670205n1 Homo sapiens cDNA clone
 L17N670205n1-46-D04 5', mRNA sequence.
 ACCESSION CB163763
 VERSION CB163763.1 GI:28149889
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 575)
 AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.U., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr

Plate: 46 row: D column: 04
High quality sequence stop: 575.

FEATURES

Location/Qualifiers

1..575
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="L17N670205n1-46-D04"
/sex="F"
/lab_host="Top10F"
/note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; The library was contributed by the Soares Laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."

ORIGIN

Alignment Scores:

Pred. No.:	3,61e-36	Length:	575
Score:	323.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x CB156803 (1-575)

QY 1 AlValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAspAspGluLeuAla 20
361 GCAGTGTGCCAGGATGGATGTACGACTACACCGCGAGATGACGATGAGCTGGCC 420
QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrPlyGly 40
421 TTCACACAGGGCCAGATCATCAACGTCCTCAACACAGAGGACCCCTGACTGTGAAAGCA 480
DB 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
481 GAAGTCATGACAGCAAGTGGGGCTCTTCCATCCATCATATATGTGAAGTGACACAGAC 537

RESULT 4
CB156803 583 bp mRNA linear EST 29-JAN-2003
LOCUS K-EST0215771 L17N670205n1 Homo sapiens cDNA clone
DEFINITION L17N670205n1-4-C07 5', mRNA sequence.
ACCESSION CB156803
VERSION CB156803.1 GI:28141931
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 583) Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE
JOURNAL 21C Frontier Korean EST Project 2001
COMMENT Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 4 row: C column: 07
High quality sequence stop: 583.

FEATURES
source 1..583
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"

/clone="L17N670205n1-4-C07"
/sex="F"
/lab_host="Top10F"
/note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; The library was contributed by the Soares Laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."

ORIGIN

Alignment Scores:

Pred. No.:	3,68e-36	Length:	583
Score:	323.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x CB156803 (1-583)

QY 1 AlValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAspAspGluLeuAla 20
361 GCAGTGTGCCAGGATGGATGTACGACTACACCGCGAGATGACGATGAGCTGGCC 420
QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrPlyGly 40
421 TTCACACAGGGCCAGATCATCAACGTCCTCAACACAGAGGACCCCTGACTGTGAAAGCA 480
DB 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
481 GAAGTCATGACAGCAAGTGGGGCTCTTCCATCCATCATATATGTGAAGTGACACAGAC 537

RESULT 5
CB215044 591 bp mRNA linear EST 06-FEB-2003
LOCUS NISC_np01906.y1 NICHD_HS_Utl Homo sapiens cDNA clone IMAGE:5936818
DEFINITION 5', mRNA sequence.
ACCESSION CB215044
VERSION CB215044.1 GI:28263236
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 591) Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
CDNA Library Preparation:
DNA Sequencing By: The I.M.A.G.E. Consortium/LLNL
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLML1362 row: M column: 11
Seq primer: M13RP reverse primer (ABI).
Location/Qualifiers

FEATURES

source 1..591

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936818"
/sex="female"
/tissue_type="normal endometrium, late proliferative
phase, cycle day 13"
/lab_host="DH10B (71-resistant)"
/clone_lib="NICHD_HS_Utl"

/note="Organ: uterus; Vector: PCMV-SPORT6.1.ccdh (ResGen, Invitrogen Corporation); Site_1: NotI; Site_2: EcoRV; Cloned unidirectionally from microquantity amounts of mRNA from normal endometrial tissue (late proliferative phase, cycle day 13). Average insert size 1.9 kb. Library constructed by ResGen (Invitrogen Corporation)."

5'-GACTAGTTCAGATCCGACGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ALIGNMENT SCORES:

Pred. No.: 3,75e-36 Length: 591
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

ALIGNMENT SCORES:

Pred. No.: 8.82e-36 Length: 1062
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x CB215044 (1-591)

US-09-720-934-2_COPY_1080_1138 (1-59) x BQ898718 (1-1062)

QY 1 ALaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20

QY 1 ALaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20

DB 153 GCACTGTGCCAGGCGATGGAGTTGGAGTTACGACTACACCGCGCAGATGACGATGAGCTGGCC 212

DB 574 GCACTGTGCCAGGCGATGGAGTTGGAGTTACGACTACACCGCGCAGATGACGATGAGCTGGCC 633

QY 21 PheAsnIysGlyGlnIleIleAsnValLeuAsnIysGluAspProAspTyrTrpIlysgly 40

QY 21 PheAsnIysGlyGlnIleIleAsnValLeuAsnIysGluAspProAspTyrTrpIlysgly 40

DB 213 TTCAACAAGAGGCCAGATCATCAAGCTCTCAACAAGAGGACCTGACTGAGTGGAAAGGA 272

DB 634 TTCAACAAGAGGCCAGATCATCAAGCTCTCAACAAGAGGACCTGACTGAGTGGAAAGGA 693

QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValIlySLeuThrThrAsp 59

QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValIlySLeuThrThrAsp 59

DB 273 GAAGTCATATGACAAAGTGGGGCTCTTCCCATCCAAATTATGTGAAGTGCACACAGAC 329

DB 694 GAAGTCATATGACAAAGTGGGGCTCTTCCCATCCAAATTATGTGAAGTGCACACAGAC 750

RESULT 6

RESULT 7

LOCUS BQ898718 1062 bp mRNA linear EST 16-AUG-2002

LOCUS BE888832 1064 bp mRNA linear EST 20-OCT-2000

DEFINITION AGENCOURT 8118759 Lupski dorsal root ganglion Homo sapiens cDNA

DEFINITION 601513752P1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3914940 5', mRNA sequence.

ACCESSION BQ898718

ACCESSION BE888832

VERSION BQ898718.1 GI:22290732

VERSION BE888832.1 GI:10345530

KEYWORDS EST.

KEYWORDS EST.

ORGANISM Homo sapiens (human)

ORGANISM Homo sapiens (human)

REFERENCE NIH-MGC http://mgs.nci.nih.gov/.

REFERENCE NIH-MGC http://mgs.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

COMMENT Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Dr. James R. Lupski

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

http://image.llnl.gov

Plate: L1AM13562 row: n column: 16

Plate: L1AM9737 row: h column: 13

High quality sequence start: 116

High quality sequence stop: 607.

Location/Qualifiers

Location/Qualifiers

1. 1062

1. 1064

/organism="Homo sapiens"

/organism="Homo sapiens"

/mol_type="mRNA"

/mol_type="mRNA"

/db_xref="taxon:9606"

/db_xref="taxon:9606"

/clone="IMAGE:6180303"

/clone="IMAGE:3914940"

/sex="male"

/sex="male"

/tissue_type="dorsal root ganglia"

/tissue_type="leiomyosarcoma"

/dev stage="adult, 36 yr"

/lab host="DH10B"

/clone_lib="Lupski dorsal root ganglion"

/clone_lib="NIH_MGC_71"

/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:

/note="Organ: uterus; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCAAGCGCTCCG-3' and

5'-TCGACCAAGCGCTCCG-3' and

Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BB888832 (1-1064)

QY 1 A1aValcysglnVal11leglyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
Db 122 GCAGTGGCCAGGAGTGGATGGATGACACTACACCGCGAGATGACGATGAGCTGGCC 181
QY 21 PheasnlysglyGln11le1leasnValleuasnlysglyuaspProaspTyrTyrlysgly 40
Db 182 TTCACAGGCGCCAGATCATCAACGCTCTCAACAGAGAGACCTGACTGGTGGAAAGGA 241
QY 41 GluValasnlyglnVal1g1yleuPheProSerAsnTyrVallylsleuthrThrAsp 59
Db 242 GAAGTCAATGACAGACAGTGGGCTCTTCCATCCATTAATGAGAGCTGACACAGAC 298

RESULT 8 363 bp mRNA linear EST 18-OCT-1995
LOCUS H66900
DEFINITION y71d07.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:210733 5' similar to SP:MYSC_ACACA P10569 MYOSIN IC HEAVY ;
mRNA sequence.
H66900
ACCESSION H66900.1 GI:1025640
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM

REFERENCE
AUTHORS

Mammalia: Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 363)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rife, J.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.,
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
8889549

COMMENT

TITLE
JOURNAL
MEDLINE
PUBMED
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estw@wustl.edu
Insert Size: 522
High quality sequence stops: 299
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 522 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 299.
Location/Qualifiers

FEATURES

source
1. 363
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3783574"
/db_xref="taxon:9606"
/clone="IMAGE:210733"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer

ORIGIN

[5' AACGGAAGAAATTAATTAAGATCTTTTCTTTTCTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

Alignment Scores:

Score: 6.97e-36 Length: 363
Pred. No.: 319.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 98.76% Indels: 0
DB: 14 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x H66900 (1-363)

QY 2 ValCysglnVal11leglyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21
Db 108 GTGTGCCAGGTGATGGATGTACACTACACCGCGAGATGACGATGAGCTGGCTTC 167
QY 22 AsnlysglyGln11le1leasnValleuasnlysglyuaspProaspTyrTyrlysgly 41
Db 168 AACAGGCGCCAGATCATCAACGCTCTCAACAGAGAGACCTGACTGGTGGAAAGGA 227
QY 42 ValasnlyglnVal1g1yleuPheProSerAsnTyrVallylsleuthrThrAsp 59
Db 228 GTCATATGACAGAGTGGGCTCTTCCATCCATTAATGAGAGCTGACACAGAC 281

RESULT 9

LOCUS BG894858 581 bp mRNA linear EST 05-JUN-2001
DEFINITION 355626 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BG894858
VERSION BG894858.1 GI:14305099
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM

REFERENCE
AUTHORS

Mammalia: Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 581)
Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
Vallet, J., Wise, T., Rohrer, G.A., Pertea, G., Sultana, R.,
Quackenbush, J., and Keefe, J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
22213789
12226715

COMMENT

TITLE
JOURNAL
MEDLINE
PUBMED
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCAGACG
Plate: 120 row: G column: 3
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers

FEATURES

source
1. 581
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1PIG"
/note="Vector: pcwv SPOR6; Site_1: NotI; Site_2: SalI;

Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

ORIGIN

Alignment Scores:
Pred. No.: 2.69e-35 Length: 581
Score: 317.00 Matches: 57
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.61% Mismatches: 0
Query Match: 98.14% Indels: 0
DB: 12 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BG894858 (1-581)

QY 1 ALaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
DB 167 GCAGTGTCCACGGCGATCGGATGACACTACACCGCGAGAAAGACAGAGAGCTGGCC 226

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTriplysGly 40
DB 227 TTCAACAAGGGCCAGGTATCAAGCTCTTCAACAAGAGAGACCCGACTGGTGGAGGGA 286

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
DB 287 GAAGTCCACGGCCAAAGTGGGGCTTCCCGTCCAACTACGTGAAGCTGACCAAGAC 343

RESULT 10
BG894850 612 bp mRNA linear EST 05-JUN-2001
LOCUS 35615 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BG894850
VERSION BG894850.1 GI:14305091
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa

REFERENCE
AUTHORS Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
Vallet,J., Wise,T., Kohner,G.A., Petrea,G., Sultana,R.,
Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Quackenbush,J. and Keele,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)

JOURNAL
MEDLINE 22213789
COMMENT 12226715
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mismatch 18
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCCACTGACGACG
Plate: 120 row: B column: 4
Seq primer: ATTGAGTGACCTATAG.
Location/Qualifiers
1..612
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="faoled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/note="Vector: pCMV SPOR6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

FEATURES

ORIGIN
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

ORIGIN

Alignment Scores:
Pred. No.: 2.9e-35 Length: 612
Score: 317.00 Matches: 57
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.61% Mismatches: 0
Query Match: 98.14% Indels: 0
DB: 12 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BG894850 (1-612)

QY 1 ALaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
DB 167 GCAGTGTCCACGGCGATCGGATGACACTACACCGCGAGAAAGACAGAGAGCTGGCC 226

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTriplysGly 40
DB 227 TTCAACAAGGGCCAGGTATCAAGCTCTTCAACAAGAGAGACCCGACTGGTGGAGGGA 286

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
DB 287 GAAGTCCACGGCCAAAGTGGGGCTTCCCGTCCAACTACGTGAAGCTGACCAAGAC 343

RESULT 11
BU321701 693 bp mRNA linear EST 28-NOV-2002
LOCUS 603486370P1 CSEQCHN62 Gallus gallus cDNA clone CHEST381118 5', mRNA
DEFINITION
ACCESSION BU321701
VERSION BU321701.1 GI:25829702
KEYWORDS EST.
SOURCE Gallus gallus
ORGANISM Gallus gallus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 693)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.D.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

JOURNAL
MEDLINE 22335534
COMMENT 12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612089930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..693
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST381118"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN62"
/note="Organ: heads; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996) 791, except that a significantly longer

FEATURES

ORIGIN
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institutes of Mental Health (NIMH), Hemm Chinn, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.:	1,316-34	Length:	694
Score:	313.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	96.61%	Mismatches:	0
Query Match:	96.90%	Indels:	0
DB:	13	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x BU054334 (1-694)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAla 20
Db 357 GCAgTGTGCCAGGTGATCGGATCGGATGACATTACACCCCGAGAACGATGAGAACTAGCC 416
QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrLysGly 40
Db 417 TTTCGCAAAAGCCAGATCATCATCGTCTCTCAACAAGAGAGACCCGAGACTGTGGAAAGGA 476

QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
Db 477 GAAGTCAGTGGCGAAGTGGGCTCTTCCATCCCAATTATGTAAAGCTGACCAAGAC 533

RESULT 14

LOCUS B0922629 782 bp mRNA linear EST 05-JUN-2001
DEFINITION 602821583F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4950524 5',
mRNA sequence.
ACCESSION B0922629
VERSION B0922629.1 GI:14303105
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 782)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: InCyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M10905 row: 8 column: 21
High quality sequence shop: 716.

FEATURES

source
1..782
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4950524"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"

ORIGIN

Alignment Scores:

Pred. No.:	1,566-34	Length:	782
Score:	313.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	96.61%	Mismatches:	0
Query Match:	96.90%	Indels:	0
DB:	12	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x B0922629 (1-782)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAla 20
Db 138 GCAgTGTGCCAGGTGATCGGATCGGATGACATTACACCCCGAGAACGATGAGAACTAGCC 197

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrLysGly 40
Db 198 TTTCGCAAAAGCCAGATCATCATCGTCTCTCAACAAGAGAGACCCGAGACTGTGGAAAGGA 257

QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
Db 258 GAAGTCAGTGGCGAAGTGGGCTCTTCCATCCCAATTATGTAAAGCTGACCAAGAC 314

RESULT 15

LOCUS B0614534 782 bp mRNA linear EST 20-FEB-2003
DEFINITION UI-M-EV0-cbg-1-22-0-UI.r1 NIH_BMAP_EV0 Mus musculus cDNA clone
UI-M-EV0-cbg-1-22-0-UI 5', mRNA sequence.
ACCESSION B0614534
VERSION B0614534.1 GI:23280749
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 782)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

source

Seq primer: pYX-5.
Location/Qualifiers
1..782
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-EV0-cbg-1-22-0-UI"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_EV0"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded

cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pUX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGGGTGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemmi Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.:	1.56e-34	Length:	782
Score:	313.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	96.61%	Mismatches:	0
Query Match:	96.90%	Indels:	0
DB:	13	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x BU614534 (1-782)

QY	1	AlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAla	20
Db	427	GCAgTGTGCCAGGTGATCGGATGTACGATTACACGCCGACGATGACGAACCTAGCC	486
QY	21	PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGlyAspProAspTyrTyrPlysGly	40
Db	487	TTcAGCAAAAGGCCAGATCATCAACGTCTTCAACAAGAGGACCCGACTGTGAAAGGA	546
QY	41	GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValIysLeuThrThraAsp	59
Db	547	GAAGTCAGTGCGCAGAGTGGGCTCTCCATCCATTATGTAAAGCTGACCAACAGAC	603

Search completed: July 1, 2004, 19:49:00
Job time : 1780.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 11:41:07 ; Search time 188.286 Seconds

(without alignments)
1376.313 Million cell updates/sec

Title: US-09-720-934-2_COPY_740_800

Perfect score: 333 1 VKVVYRYALYPFBSRSHDEI.....GGELKKGKGWFPANVAKXP 61

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2.1/USPTO.spool_p/US09720934/rnat.30062004_064539_13422/app_query.fasta_1.1386
-DB=N Geneseg 29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934 @CGN 1.1 1596 @rnat.30062004_064539_13422 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseg 29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	333	100.0	1389	2	AAT39795 Human SH3
2	333	100.0	2874	5	AAS84762 DNA encod
3	333	100.0	3231	2	AAZ34574 Human SH3
4	333	100.0	3319	4	ABK43498 Human SH3
5	333	100.0	3466	4	AAI63825 Human SH3
6	333	100.0	3723	3	AAZ39009 Human SH3
7	333	100.0	5082	3	AAZ39008 Human SH3
8	333	100.0	5144	3	AAZ39025 Mouse ESE

9	333	100.0	5195	2	AAZ34572 Human SH3
10	333	100.0	5199	2	AAZ34570 Human SH3
11	333	100.0	5738	3	AAZ39024 Mouse ESE
12	320.5	96.2	5458	2	AAZ34571 Human SH3
13	320.5	96.2	7435	5	AAS84763 DNA encod
14	216	64.9	2873	2	AAT39799 Human SH3
15	216	64.9	3593	3	AAZ39011 Human SH3
16	216	64.9	3746	4	ABK43586 DNA encod
17	216	64.9	4210	4	AAS27090 Human SH3
18	216	64.9	4210	9	ADB93268 Human SH3
19	216	64.9	4447	4	AAS02055 DNA encod
20	216	64.9	4625	3	AAZ39010 Human SH3
21	216	64.9	4975	3	AAZ39027 Human SH3
22	216	64.9	5828	6	AAI47247 Allergic
23	216	64.9	6014	3	AAZ39026 Mouse ESE
24	216	64.9	6103	4	AAK52332 Human SH3
25	184	55.3	3981	4	ABL01995 Human SH3
26	184	55.3	7235	4	ABL01994 Drosophila
27	182	54.7	270	4	AAI19657 Human SH3
28	182	54.7	270	4	ABA64683 Human SH3
29	182	54.7	270	4	AAI44852 Human SH3
30	182	54.7	270	4	ABA46803 Human SH3
31	182	54.7	270	4	ABA31807 Human SH3
32	182	54.7	270	4	AAK38849 Human SH3
33	182	54.7	270	4	AAK31319 Human SH3
34	182	54.7	270	4	ABS38434 Human SH3
35	182	54.7	270	5	AAI05376 Human SH3
36	182	54.7	270	6	ABS12930 Human SH3
37	182	54.7	286	4	AAI23124 Human SH3
38	182	54.7	286	4	ABA68217 Human SH3
39	182	54.7	286	4	AAI48432 Human SH3
40	182	54.7	286	4	ABA50277 Human SH3
41	182	54.7	286	4	AAI48432 Human SH3
42	182	54.7	286	4	AAK42353 Human SH3
43	182	54.7	286	4	AAK416592 Human SH3
44	182	54.7	286	5	AAI08779 Human SH3
45	182	54.7	286	6	ABS16408 Human SH3

ALIGNMENTS

RESULT 1	AAZ39795	Standard, DNA, 1389 BP.
XX	AAZ39795	
AC	AAZ39795	
XX	AAZ39795	
DT	19-FEB-1998	(first entry)
XX	19-FEB-1998	
DE	Human SH3P17 gene.	
XX	Human SH3P17 gene.	
KW	Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;	
KW	cellular signalling element; cellular structural element; malignancy;	
KW	protein identification; functional domain; protein screening;	
KW	cellular signal transduction process; ss.	
OS	Homo sapiens.	
XX	Homo sapiens.	
XX	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	CDS	1..1389
FT	FT	/tag= a
FT	FT	/product= "SH3P17"
XX	XX	
PN	W09631625-A1.	
XX	W09631625-A1.	
PD	10-OCT-1996.	
XX	10-OCT-1996.	
PE	04-APR-1996;	96MO-US004454.
XX	04-APR-1996;	
PR	07-APR-1995;	95US-00417872.
PR	03-APR-1996;	96US-00630915.
XX	03-APR-1996;	
PA	(CYTO-) CYTOGEN CORP.	

XX	(UYNC-) UNIV NORTH CAROLINA.	
XX	Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;	
XX	WPI; 1996-465045/46.	
DR	P-PSDB; AAM05395.	
XX		
PT	Identifying polypeptide(s) having specific functional domain (esp. SH3	
PT	domain) - comprises detecting selective binding to recognition unit,	
PT	regardless of sequence homology.	
XX		
PS	Claim 76; Fig 50; 174pp; English.	
XX		
CC	AAI39786-R39803 represent novel human and mouse genes encoding Src-	
CC	homology region 3 (SH3) domain containing proteins that can be used in	
CC	the method of the invention. SH3 domain containing proteins play a role	
CC	in signalling and structural elements of cells. The method of the	
CC	invention is for identifying polypeptides containing functional domains	
CC	of interest (especially SH3 domains). The method comprises contacting a	
CC	multivalent recognition unit (RU) complex with a number of peptides and	
CC	identifying polypeptides having a selective binding affinity for the RU	
CC	complex. The method is based on functional similarities and does not rely	
CC	on sequence similarities. Prior methods only gave limited success for	
CC	identifying proteins which contain an SH3 domain due to the minimal	
CC	sequence homology among known SH3 proteins. It has been found that small	
CC	peptide RUs in multivalent form have reduced specificity for a given	
CC	functional domain compared to monomer RUs. Multivalent RU complexes are	
CC	particularly suited to screening for polypeptides containing functional	
CC	domains that are similar to, but not identical in sequence to, the	
CC	original target functional domain. The new method enables proteins having	
CC	a common function to be identified. Identification of novel SH3 proteins	
CC	will be useful for a better understanding of cell growth, malignancy,	
CC	signal transduction processes, etc. New candidate drugs can be	
CC	identified, and their specificities (e.g. pharmacological activities) can	
CC	be assessed using the method of the invention	
XX		
SQ	Sequence 1389 BP; 404 A; 370 C; 357 G; 258 T; 0 U; 0 Other;	
	Alignment Scores:	
	Pred. No.: 6,9e-35 Length: 1389	
	Score: 333.00 Matches: 61	
	Percent Similarity: 100.00% Conservative: 0	
	Best Local Similarity: 100.00% Mismatches: 0	
	Query Match: 100.00% Indels: 0	
	DB: 2 Gaps: 0	
	US-09-720-934-2_COPY_740_800 (1-61) x AAT39795 (1-1389)	
QY	1 VALVSVAVAlTYrTYrArngAlaLeuTYrPPoPhGluSerArGerHisAspGluIle 20	
Db	172 GAAAGAGTGATTAACCGGAGACGTACCCCTTGATTCGAAAGCCATGTGAATC 231	
QY	21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrngIyGlnProGlyTIP 40	
Db	232 ACATATCAGCGAGAACATATGCAAGTGTGATGAAGCAAACTGAGAAACCGGCTGG 291	
QY	41 LeuGlyGlyGluLeuIySgIyIySThngIYTIPheProAlaSerTYrAlaGluIyIle 60	
Db	292 CTGGAGAGGAATTAAGAAAGAAAGACAGGGGTGTTCCCTGCACAACTATGACAGGAAATC 351	
QY	61 Pro 61	
Db	352 CCA 354	
	RESULT 2	
	AA584762	
ID	AA584762 standard; cDNA; 2874 BP.	
XX	AA584762;	
AC		
XX	13-FEB-2002 (first entry)	
DT		
XX		
DE	DNA encoding novel human diagnostic protein #20566.	

XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	
XX	Homo sapiens.
OS	
XX	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
XX	30-MAR-2001; 2001WO-US006631.
PF	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
XX	
XX	(HXSE-) HXSEQ INC.
PA	
XX	
PI	Dymanac RT, Liu C, Tang YT;
DR	WPI; 2001-639362/73.
DR	P-PSDB; ABG20575.
XX	
XX	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 1; SEQ ID NO 20566; 103bp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (II) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostic, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 2874 BP; 940 A; 679 C; 703 G; 552 T; 0 U; 0 Other;
US-09-720-934-2_COPY_740_800 (1-61) x AAS84762 (1-2874)	
Alignment Scores:	
Pred. No.:	1.7e-34 Length: 2874
Score:	333.00 Matches: 61
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	5 Gaps: 0
QY	1 ValysValValTyrTyrArgAlaLeuYrPpPheGluSerArgSerHisAspGluIle 20
Db	913 GTAAAGGTGTGTATTATCCGGGCACTGTHCCCTTTGAATCAGAAGCAAGATGAATC 972
QY	21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGlnProGlyTyr 40
Db	973 ACGATCAAGCGACGAGACATAGTCATGTGGATGAAGCAAACTGGAGAACCCGGCTGG 1032
QY	41 LeuGlyGlyGlnLeuLysGlyLysThrGlyTTPpPheProAlaAsnTyrAlaGlnLysIle 60

Db 1033 CTTGAGGAGAAATTAAAGAAAGACAGGCTGTTCCCTCGCAACTATGACAGAAATTC 1092

QY 61 Pro 61
|||
Db 1093 CCA 1095

RESULT 3
AAZ34574
ID AAZ34574 standard; cDNA, 3231 BP.

XX AC AAZ34574;
XX DT 01-FEB-2000 (first entry)
XX DE Human SH3D1A cDNA clone 9.
XX

SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
OS Homo sapiens.
XX

XX Key Location/Qualifiers
FT CDS 2..1927
FT /*tag= a
XX

XX MO9953062-A2.
XX 21-OCT-1999.
XX 16-APR-1999; 99MO-US008371.
XX 16-APR-1998; 98US-0082007P.
XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
XX Korenberg JR, Chen X;
XX MPI: 1999-633829/54.
XX P-PSDB; AAY32158.
XX

Nucleic acid from the human SH3D1A gene and its products, useful for the
diagnosis and treatment of myeloproliferative disorders and leukemia.
Claim 2; Fig 14; 99p; English.

This is the nucleotide sequence of a non-full-length cDNA (clone 9)
corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
contributes to the development of platelets and the pathogenesis of
leukaemias, both in general and in particular those involving the
megakaryocytic lineage. The SH3D1A gene maps to the small candidate
region for low platelets on chromosome 21. Sequencing of 5 different
sizes of cDNA clone from foetal brain (see AAZ34570-74) suggests that at
least 3 isoforms exist. The invention provides methods for the diagnosis
and treatment of megakaryocytic abnormality, myeloproliferative disorder,
platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
association of gains in chromosome 21, low platelets in deletion for 21,
abnormalities, dysfunctions and disorders including brain malformations
and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
and colpocephaly. Methods are also provided for: suppressing cells unable
to regulate themselves; screening for a somatic alteration in the SH3D1A
gene; monitoring the progress and adequacy of a treatment; monitoring
tumour risk progress or megakaryocytic abnormality; myeloproliferative
disorder, haematopoietic disorder, platelet disorder or leukaemia; and
treatment of a subject (including a prenatal subject) having
megakaryocytic abnormality, myeloproliferative disorder, platelet
disorder, leukaemia or neural disorder using a nucleic acid that
expresses SH3D1A or its antisense nucleic acid

Sequence 3231 BP, 1004 A; 721 C; 712 G; 794 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.97e-34 Length: 3231
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AAZ34574 (1-3231)

QY 1 ValysValValYTYrTgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
Db 710 GTAAAGTGTGATATACCGGCGACCTGTACCCCTTGAATCCAGAACCATGATGAATTC 769

QY 21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGlnProGlyTyr 40
Db 770 ACTATCCAGCCAGAGACATGATGCTAGTGGATGATGAAGCCAAACTGGAGAAACCGGCTGG 829

QY 41 LeuGlyGlyGluLeuValysGlyTyrThrGlyTyrPheProAlaAenTYrAlaGluIle 60
Db 830 CTTGAGGAGAAATTAAAGAAAGACAGGCTGTTCCCTCGCAACTATGACAGAAATTC 889

QY 61 Pro 61
|||
Db 890 CCA 892

RESULT 4
ID ABK43498
XX ABK43498 standard; cDNA, 3319 BP.
XX
XX AC ABK43498;
XX
XX DT 05-JUN-2002 (first entry)
XX

DE DNA encoding novel central nervous system protein #78.

KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.

OS Homo sapiens.
XX
XX WO20015318-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001332.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190073P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234222P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246533P.
PR 08-NOV-2000; 2000US-0246534P.
PR 08-NOV-2000; 2000US-0246535P.
PR 08-NOV-2000; 2000US-0246536P.
PR 08-NOV-2000; 2000US-0246537P.
PR 08-NOV-2000; 2000US-0246538P.
PR 08-NOV-2000; 2000US-0246539P.
PR 08-NOV-2000; 2000US-0246540P.
PR 08-NOV-2000; 2000US-0246541P.
PR 08-NOV-2000; 2000US-0246542P.
PR 08-NOV-2000; 2000US-0246543P.
PR 08-NOV-2000; 2000US-0246544P.
PR 08-NOV-2000; 2000US-0246545P.
PR 08-NOV-2000; 2000US-0246546P.
PR 08-NOV-2000; 2000US-0246547P.
PR 08-NOV-2000; 2000US-0246548P.
PR 08-NOV-2000; 2000US-0246549P.
PR 08-NOV-2000; 2000US-0246550P.
PR 08-NOV-2000; 2000US-0246551P.
PR 08-NOV-2000; 2000US-0246552P.
PR 08-NOV-2000; 2000US-0246553P.
PR 08-NOV-2000; 2000US-0246554P.
PR 08-NOV-2000; 2000US-0246555P.
PR 08-NOV-2000; 2000US-0246556P.
PR 08-NOV-2000; 2000US-0246557P.
PR 08-NOV-2000; 2000US-0246558P.
PR 08-NOV-2000; 2000US-0246559P.
PR 08-NOV-2000; 2000US-0246560P.
PR 08-NOV-2000; 2000US-0246561P.
PR 08-NOV-2000; 2000US-0246562P.
PR 08-NOV-2000; 2000US-0246563P.
PR 08-NOV-2000; 2000US-0246564P.
PR 08-NOV-2000; 2000US-0246565P.
PR 08-NOV-2000; 2000US-0246566P.
PR 08-NOV-2000; 2000US-0246567P.
PR 08-NOV-2000; 2000US-0246568P.
PR 08-NOV-2000; 2000US-0246569P.
PR 08-NOV-2000; 2000US-0246570P.
PR 08-NOV-2000; 2000US-0246571P.
PR 08-NOV-2000; 2000US-0246572P.
PR 08-NOV-2000; 2000US-0246573P.
PR 08-NOV-2000; 2000US-0246574P.
PR 08-NOV-2000; 2000US-0246575P.
PR 08-NOV-2000; 2000US-0246576P.
PR 08-NOV-2000; 2000US-0246577P.
PR 08-NOV-2000; 2000US-0246578P.
PR 08-NOV-2000; 2000US-0246579P.
PR 08-NOV-2000; 2000US-0246580P.
PR 08-NOV-2000; 2000US-0246581P.
PR 08-NOV-2000; 2000US-0246582P.
PR 08-NOV-2000; 2000US-0246583P.
PR 08-NOV-2000; 2000US-0246584P.
PR 08-NOV-2000; 2000US-0246585P.
PR 08-NOV-2000; 2000US-0246586P.
PR 08-NOV-2000; 2000US-0246587P.
PR 08-NOV-2000; 2000US-0246588P.
PR 08-NOV-2000; 2000US-0246589P.
PR 08-NOV-2000; 2000US-0246590P.
PR 08-NOV-2000; 2000US-0246591P.
PR 08-NOV-2000; 2000US-0246592P.
PR 08-NOV-2000; 2000US-0246593P.
PR 08-NOV-2000; 2000US-0246594P.
PR 08-NOV-2000; 2000US-0246595P.
PR 08-NOV-2000; 2000US-0246596P.
PR 08-NOV-2000; 2000US-0246597P.
PR 08-NOV-2000; 2000US-0246598P.
PR 08-NOV-2000; 2000US-0246599P.
PR 08-NOV-2000; 2000US-0246600P.
PR 08-NOV-2000; 2000US-0246601P.
PR 08-NOV-2000; 2000US-0246602P.
PR 08-NOV-2000; 2000US-0246603P.
PR 08-NOV-2000; 2000US-0246604P.
PR 08-NOV-2000; 2000US-0246605P.
PR 08-NOV-2000; 2000US-0246606P.
PR 08-NOV-2000; 2000US-0246607P.
PR 08-NOV-2000; 2000US-0246608P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246612P.
PR 08-NOV-2000; 2000US-0246613P.
PR 08-NOV-2000; 2000US-0246614P.
PR 08-NOV-2000; 2000US-0246615P.
PR 08-NOV-2000; 2000US-0246616P.
PR 08-NOV-2000; 2000US-0246617P.
PR 08-NOV-2000; 2000US-0246618P.
PR 08-NOV-2000; 2000US-0246619P.
PR 08-NOV-2000; 2000US-0246620P.
PR 08-NOV-2000; 2000US-0246621P.
PR 08-NOV-2000; 2000US-0246622P.
PR 08-NOV-2000; 2000US-0246623P.
PR 08-NOV-2000; 2000US-0246624P.
PR 08-NOV-2000; 2000US-0246625P.
PR 08-NOV-2000; 2000US-0246626P.
PR 08-NOV-2000; 2000US-0246627P.
PR 08-NOV-2000; 2000US-0246628P.
PR 08-NOV-2000; 2000US-0246629P.
PR 08-NOV-2000; 2000US-0246630P.
PR 08-NOV-2000; 2000US-0246631P.
PR 08-NOV-2000; 2000US-0246632P.
PR 08-NOV-2000; 2000US-0246633P.
PR 08-NOV-2000; 2000US-0246634P.
PR 08-NOV-2000; 2000US-0246635P.
PR 08-NOV-2000; 2000US-0246636P.
PR 08-NOV-2000; 2000US-0246637P.
PR 08-NOV-2000; 2000US-0246638P.
PR 08-NOV-2000; 2000US-0246639P.
PR 08-NOV-2000; 2000US-0246640P.
PR 08-NOV-2000; 2000US-0246641P.
PR 08-NOV-2000; 2000US-0246642P.
PR 08-NOV-2000; 2000US-0246643P.
PR 08-NOV-2000; 2000US-0246644P.
PR 08-NOV-2000; 2000US-0246645P.
PR 08-NOV-2000; 2000US-0246646P.
PR 08-NOV-2000; 2000US-0246647P.
PR 08-NOV-2000; 2000US-0246648P.
PR 08-NOV-2000; 2000US-0246649P.
PR 08-NOV-2000; 2000US-0246650P.
PR 08-NOV-2000; 2000US-0246651P.
PR 08-NOV-2000; 2000US-0246652P.
PR 08-NOV-2000; 2000US-0246653P.
PR 08-NOV-2000; 2000US-0246654P.
PR 08-NOV-2000; 2000US-0246655P.
PR 08-NOV-2000; 2000US-0246656P.
PR 08-NOV-2000; 2000US-0246657P.
PR 08-NOV-2000; 2000US-0246658P.
PR 08-NOV-2000; 2000US-0246659P.
PR 08-NOV-2000; 2000US-0246660P.
PR 08-NOV-2000; 2000US-0246661P.
PR 08-NOV-2000; 2000US-0246662P.
PR 08-NOV-2000; 2000US-0246663P.
PR 08-NOV-2000; 2000US-0246664P.
PR 08-NOV-2000; 2000US-0246665P.
PR 08-NOV-2000; 2000US-0246666P.
PR 08-NOV-2000; 2000US-0246667P.
PR 08-NOV-2000; 2000US-0246668P.
PR 08-NOV-2000; 2000US-0246669P.
PR 08-NOV-2000; 2000US-0246670P.
PR 08-NOV-2000; 2000US-0246671P.
PR 08-NOV-2000; 2000US-0246672P.
PR 08-NOV-2000; 2000US-0246673P.
PR 08-NOV-2000; 2000US-0246674P.
PR 08-NOV-2000; 2000US-0246675P.
PR 08-NOV-2000; 2000US-0246676P.
PR 08-NOV-2000; 2000US-0246677P.
PR 08-NOV-2000; 2000US-0246678P.
PR 08-NOV-2000; 2000US-0246679P.
PR 08-NOV-2000; 2000US-0246680P.
PR 08-NOV-2000; 2000US-0246681P.
PR 08-NOV-2000; 2000US-0246682P.
PR 08-NOV-2000; 2000US-0246683P.
PR 08-NOV-2000; 2000US-0246684P.
PR 08-NOV-2000; 2000US-0246685P.
PR 08-NOV-2000; 2000US-0246686P.
PR 08-NOV-2000; 2000US-0246687P.
PR 08-NOV-2000; 2000US-0246688P.
PR 08-NOV-2000; 2000US-0246689P.
PR 08-NOV-2000; 2000US-0246690P.
PR 08-NOV-2000; 2000US-0246691P.
PR 08-NOV-2000; 2000US-0246692P.
PR 08-NOV-2000; 2000US-0246693P.
PR 08-NOV-2000; 2000US-0246694P.
PR 08-NOV-2000; 2000US-0246695P.
PR 08-NOV-2000; 2000US-0246696P.
PR 08-NOV-2000; 2000US-0246697P.
PR 08-NOV-2000; 2000US-0246698P.
PR 08-NOV-2000; 2000US-0246699P.
PR 08-NOV-2000; 2000US-0246700P.
PR 08-NOV-2000; 2000US-0246701P.
PR 08-NOV-2000; 2000US-0246702P.
PR 08-NOV-2000; 2000US-0246703P.
PR 08-NOV-2000; 2000US-0246704P.
PR 08-NOV-2000; 2000US-0246705P.
PR 08-NOV-2000; 2000US-0246706P.
PR 08-NOV-2000; 2000US-0246707P.
PR 08-NOV-2000; 2000US-0246708P.
PR 08-NOV-2000; 2000US-0246709P.
PR 08-NOV-2000; 2000US-0246710P.
PR 08-NOV-2000; 2000US-0246711P.
PR 08-NOV-2000; 2000US-0246712P.
PR 08-NOV-2000; 2000US-0246713P.
PR 08-NOV-2000; 2000US-0246714P.
PR 08-NOV-2000; 2000US-0246715P.
PR 08-NOV-2000; 2000US-0246716P.
PR 08-NOV-2000; 2000US-0246717P.
PR 08-NOV-2000; 2000US-0246718P.
PR 08-NOV-2000; 2000US-0246719P.
PR 08-NOV-2000; 2000US-0246720P.
PR 08-NOV-2000; 2000US-0246721P.
PR 08-NOV-2000; 2000US-0246722P.
PR 08-NOV-2000; 2000US-0246723P.
PR 08-NOV-2000; 2000US-0246724P.
PR 08-NOV-2000; 2000US-0246725P.
PR 08-NOV-2000; 2000US-0246726P.
PR 08-NOV-2000; 2000US-0246727P.
PR 08-NOV-2000; 2000US-0246728P.
PR 08-NOV-2000; 2000US-0246729P.
PR 08-NOV-2000; 2000US-0246730P.
PR 08-NOV-2000; 2000US-0246731P.
PR 08-NOV-2000; 2000US-0246732P.
PR 08-NOV-2000; 2000US-0246733P.
PR 08-NOV-2000; 2000US-0246734P.
PR 08-NOV-2000; 2000US-0246735P.
PR 08-NOV-2000; 2000US-0246736P.
PR 08-NOV-2000; 2000US-0246737P.
PR 08-NOV-2000; 2000US-0246738P.
PR 08-NOV-2000; 2000US-0246739P.
PR 08-NOV-2000; 2000US-0246740P.
PR 08-NOV-2000; 2000US-0246741P.
PR 08-NOV-2000; 2000US-0246742P.
PR 08-NOV-2000; 2000US-0246743P.
PR 08-NOV-2000; 2000US-0246744P.
PR 08-NOV-2000; 2000US-0246745P.
PR 08-NOV-2000; 2000US-0246746P.
PR 08-NOV-2000; 2000US-0246747P.
PR 08-NOV-2000; 2000US-0246748P.
PR 08-NOV-2000; 2000US-0246749P.
PR 08-NOV-2000; 2000US-0246750P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-581633/65.

P-PSDB; AAU87168.

New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 1; SEQ ID NO 88; 837pp; English.

The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (I) and polypeptides (III) encoded by (I), are used to treat a medical conditions and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. Rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, atherosclerosis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes

CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Alignment Scores:

Pred. NO.:	2.04e-34	Length:	3319
Score:	333.00	Matches:	61
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x ABR43498 (1-3319)

QY 1 VallystValValYTYTYTARGAlaLeuTYRProPheGluSerArgSerHisAspGluIle 20

Db 2388 GMAAAAGTGGTGTATTAACCGGCACTGTACCCCTTGATCCAGAACCATGATGAATC 2447

QY 21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40

Db 2448 ACATATCCAGCCAGAGACATAGTCATGTGTGATGAAGCCCAACTGGAGAACCCGGCTGG 2507

QY 41 LeuGlyGlyGluLeuLeuGlyGlyGlyTyrPheProAlaAsnTyrAlaGluIle 60

Db 2508 CTGGAGAGAGATTAAAGAAAGACAGGGGTGTTCTCTCCATCAACTGTGACAGAAATATC 2567

QY 61 Pro 61

Db 2568 CCA 2570

RESULT 5
ID AAI63825 standard; cDNA; 3466 BP.

XX AAI63825;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 33.

XX Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
XX cerebroprotective; nootropic; neuroprotective; antibacterial; vituicide;
XX fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
XX neuroprotective; antiallergic; hepatotropic; antidiabetic;
XX antiinflammatory; antitumor; anticonvulsant; antibacterial;
XX antiparasitic; cardiant; gene therapy; cancer; immune disorder;
XX cardiovascular disorder; neurological disease; infection; human; ss.

XX Homo sapiens.

XX WO20015308-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001309.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214865P.

PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216680P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 30-AUG-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.

CC	bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
CC	urogenital; (b) immune disorders e.g. Addison's disease, allergies,
CC	autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC	Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC	colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
CC	wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC	and parasitic infections. Note: The sequence data for this patent did not
CC	form part of the printed specification, but was obtained in electronic
CC	format directly from WIPO at ftp.wipo.int/publ/published_pct_sequences
XX	
XX	Sequence 3466 BP; 1056 A; 829 C; 861 G; 710 T; 0 U; 10 Other;
XX	
XX	Alignment Scores:
XX	Pred. No.: 2.15e-34 Length: 3466
XX	Score: 333.00 Matches: 61
XX	Percent Similarity: 100.00% Conservative: 0
XX	Best Local Similarity: 100.00% Mismatches: 0
XX	Query Match: 100.00% Indels: 0
XX	DB: 4 Gaps: 0
XX	
XX	US-09-720-934-2_COPY 740_800 (1-61) x AA153825 (1-3466)
QY	1 VallyysValValIYYTYTYrYrGAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
Db	2506 GTAAAGGCGGTGATNTACCGGGCACTGTACCCCTTGAATCCAGAAGCCATGATATAATC 2565
QY	21 ThrIleGlnPProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrp 40
Db	2566 ACTATCCAGCCAGAGGACATAGTCATGTGTGATGAAGCCAAACTGGAGAACCCCGCTGG 2622
QY	41 LeuGlyGlyGlyIleLeuLeuGlyLysThrGlyTrpPheProAlaAsnTyAlaGluIysIle 60
Db	2626 CTTGAGAGGAGATTAAAGAAAGACAGAGGTCGTTCCCTGCAGAACTATATCAGAGAAATC 2684
QY	61 Pro 61
Db	2686 CCA 2688
XX	
XX	RESULT 6
XX	AAZ39009
XX	ID AAZ39009 standard; cDNA; 3723 BP.
XX	
XX	AAZ39009;
AC	
XX	28-FEB-2000 (first entry)
DT	
XX	
DE	Mouse Esel coding sequence.
XX	
KW	Mouse; murine; Esel; Esel; endocytosis; vesicular trafficking;
KW	regulation; actin cytoskeleton; detection; cancer; infection;
KW	EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW	antiproliferative; antiviral; ss.
XX	
OS	Mus sp.
XX	
XX	WO955728-A2.
PN	
XX	04-NOV-1999.
PD	
XX	
PF	27-APR-1999; 99WO-CA000375.
XX	
PR	27-APR-1998; 98CA-02230201.
XX	
PR	05-FEB-1999; 99US-0118739P.
XX	
PA	(HSCR-) HSC RES & DEV LP.
XX	
PI	Egan SE, Wang W, Sengar A;
XX	
DR	WPI, 2000-052802/04.
XX	
DR	P-FSDB; AA157444.
XX	
PT	New nucleic acid encoding Esel and 2 proteins, involved in regulation of

PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS Claim 6; Page 40-42; 99pp; English.
XX
CC The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp1 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1 complex, then binding dynam to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,35e-34 Length: 3723
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AA239009 (1-3723)
QY 1 ValLysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
DB 2212 GTAAAGTGTATATTATCCAGCGCTGTACCCCTTGAATCCAGAGTCAGATGAGATC 2271
QY 21 ThrILeGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrP 40
DB 2272 ACCATCCAGCCAGAGATATATGTCATGTCGATGTAAGCAAGCCAGCTGAGAGCCAGGATGG 2331
QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
DB 2332 CTTGGAGGAGAGCTGAAGGGAAGACGGGATGTTCCCTGCAAACTATGAGAAAGATT 2391
QY 61 Pro 61
DB 2392 CCA 2394

RESULT 7
AA239008
ID AA239008 standard; cDNA; 5082 BP.
XX
AC AA239008;
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse Ese1 full length cDNA sequence.
XX
KW Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
PN WO955728-A2.
XX

PD 04-NOV-1999.
XX
XX
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
XX WPI: 2000-052802/04.
DR P-PSDB; AA57444.
XX
XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
PS Claim 6; Page 38-40; 99pp; English.
XX
XX The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp1 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1 complex, then binding dynam to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.46e-34 Length: 5082
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AA239008 (1-5082)
QY 1 ValLysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
DB 2470 GTAAAGTGTATATTATCCAGCGCTGTACCCCTTGAATCCAGAGTCAGATGAGATC 2529
QY 21 ThrILeGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrP 40
DB 2530 ACCATCCAGCCAGAGATATATGTCATGTCGATGTAAGCAAGCCAGCTGAGAGCCAGGATGG 2589
QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
DB 2590 CTTGGAGGAGAGCTGAAGGGAAGACGGGATGTTCCCTGCAAACTATGAGAAAGATT 2649
QY 61 Pro 61
DB 2650 CCA 2652

RESULT 8
AA239025
ID AA239025 standard; cDNA; 5144 BP.
XX

AC AA239025;
 XX 28-FEB-2000 (first entry)
 XX Mouse Ese1L coding sequence.
 DE Mouse Ese1L coding sequence.
 XX Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
 KW regulation; actin cytoskeleton; detection; cancer; infection;
 KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
 KW antiproliferative; antiviral; ss.
 XX
 OS Mus sp.
 PN WO995728-A2.
 XX 04-NOV-1999.
 PD
 XX 27-APR-1999; 99WO-CA000375.
 PF
 XX 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 XX
 PI Egan SE, Wang W, Sengar A;
 XX WPI, 2000-052802/04.
 DR P-PSDB; AAY57449.
 XX
 PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 PS
 XX Claim 6; Page 59-62; 99pp; English.
 XX
 CC The present invention specifically describes mammalian Ese1 and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Esp15 protein); vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC sequences antisense to the (I) polynucleotide; agents that downregulate
 CC expression of Ese genes or antagonists of an Ese binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Ese1 is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) or Ab are used to promote endocytosis of selected cells.
 CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Ese mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Ese-Esp15 complex, then binding dynamin to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The
 CC present sequence represents mouse Ese1L coding sequence
 CC
 XX
 SQ Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3.51e-34 Length: 5144
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_740_800 (1-61) X AA239025 (1-5144)
 QY 1 VALVSVAlValTYTrrgAlaleuTYrrPhphGluserArGserHisaspGuile 20
 DB 2212 GTAAAGGTGGTATATACCGAGCGCTGTACCCCTTGAAATCCAGAGTACGATGATC 2271
 QY 21 ThrileGlnPrGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40

DB 2272 ACCATCCAGCCAGAGATATATAGCATGTGATGAAAGCCAGACTGAGGCCAGGATGG 2331
 QY 41 LengIyGlyGluLeuIeulysGlyIyTrrnglyTrrPhphProhIasntYrAlaglLysle 60
 DB 2332 CTGGAGAGAGAGCTGAAAGGAAAGACGGGATGGTTCCCTCGAATATGACGAAAGAGATT 2331
 QY 61 Pro 61
 DB 2392 CCA 2394
 RESULT 9
 AA234572
 ID AA234572 standard; cDNA; 5195 BP.
 XX
 AC AA234572;
 XX 01-FEB-2000 (first entry)
 DT
 XX
 DE Human SH3D1A cDNA clone 11.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 239..386
 FT /*tag= a
 XX
 PN WO9953062-A2.
 XX
 XX 21-OCT-1999.
 PD
 XX 16-APR-1999; 99WO-US008371.
 PF
 XX 16-APR-1998; 98US-0082007P.
 PR
 XX (CEDA-) CEDARS STINAI HEALTH SYSTEM.
 PA
 PI Korenberg JR, Chen X;
 XX WPI: 1999-633829/54.
 DR P-PSDB; AAY32156.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 10; 99pp; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having

CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX

SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3,566-34	Length:	5195
Score:	333.00	Matches:	61
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x AAZ34570 (1-5195)

QY 1 VallysValValYTYTYrArgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
 DB 2456 GTAAAGGTGGTGTATACCGGGCACTGTACCCCTTGATCCAGAGCCATGATGAATC 2515

QY 21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
 DB 2516 ACTTCCAGCCAGAGACATGTGATGTGGATGAAGCCAACTGGAGAACCCGGCTGG 2575

QY 41 LeuGlyGlyGluLeuLeuTySGlyTySThrGlyTyrPheProAlaAsnTYrAlaGluIle 60
 DB 2576 CTTGAGAGAGAAATTAAAGAAAGACAGAGGTGTTCCCTCGAAACTATGACAGAGAAATC 2635

QY 61 Pro 61
 DB 2636 CCA 2638

RESULT 10

AAZ34570
 ID AAZ34570 standard; cDNA; 5199 BP.

AC AAZ34570;

DT 01-FEB-2000 (first entry)

XX Human SH3D1A cDNA.

XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 OS Homo sapiens.

Key	Location/Qualifiers
FT CDS	208..3642
FT	/*tag= a

PN WO953062-A2.

XX 21-OCT-1999.

PF 16-APR-1999; 99WO-US008371.

XX 16-APR-1998; 98US-0082007P.

XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.

XX Korenberg JR, Chen X;

XX WPI, 1999-633829/54.

XX P-PSDB; AAY32154.

XX Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX

PS Claim 2; Fig 5; 99pp; English.

XX

This is the nucleotide sequence of full-length cDNA corresponding to a
 CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the
 CC development of platelets and the pathogenesis of leukaemias, both in
 CC general and in particular those involving the megakaryocytic lineage. The
 CC SH3D1A gene maps to the small candidate region for low platelets on
 CC chromosome 21. Sequencing of 5 different sizes of cDNA clone (see
 CC AAZ34570-74) suggests that at least 3 isoforms exist. The invention
 CC provides methods for the diagnosis and treatment of megakaryocytic
 CC abnormality, myeloproliferative disorder, platelet disorder, acute
 CC leukaemia, neural disorders, thrombocytopenia, platelet disorder on
 CC chromosome 21, low platelets in deletion for 21, association of gains in
 CC chromosome 21 with leukaemias, neural abnormalities, dysfunctions and
 CC disorders including brain malformations, and corresponding cognitive
 CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
 CC also provided for: suppressing cells unable to regulate themselves;
 CC screening for a somatic alteration in the SH3D1A gene; monitoring the
 CC progress and adequacy of a treatment; monitoring tumour risk progress or
 CC megakaryocytic abnormality, myeloproliferative disorder, haematopoietic
 CC disorder, platelet disorder or leukaemia; and treatment of a subject
 CC (including a prenatal subject) having megakaryocytic abnormality,
 CC myeloproliferative disorder, platelet disorder, leukaemia or neural
 CC disorder using a nucleic acid that expresses SH3D1A or its antisense
 CC nucleic acid
 XX

SQ Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3,566-34	Length:	5199
Score:	333.00	Matches:	61
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x AAZ34570 (1-5199)

QY 1 VallysValValYTYTYrArgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
 DB 2425 GTAAAGGTGGTGTATACCGGGCACTGTACCCCTTGATCCAGAGCCATGATGAATC 2484

QY 21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
 DB 2485 ACTATCCAGCCAGAGACATGTGATGTGGATGAAGCCAACTGGAGAACCCGGCTGG 2544

QY 41 LeuGlyGlyGluLeuLeuTySGlyTySThrGlyTyrPheProAlaAsnTYrAlaGluIle 60
 DB 2545 CTTGAGAGAGAAATTAAAGAAAGACAGAGGTGTTCCCTCGAAACTATGACAGAGAAATC 2604

QY 61 Pro 61

DB 2605 CCA 2607

RESULT 11

AAZ39024
 ID AAZ39024 standard; cDNA; 5738 BP.

AC AAZ39024;

DT 28-FEB-2000 (first entry)

XX Mouse Escll cDNA sequence.

XX Mouse; murine; Escl; Escl; endocytosis; vesicular trafficking;
 KW regulation; actin cytoskeleton; detection; cancer; infection;
 KW EH-domain and SH3 domain regulator of endocytosis; anticancer;
 KW antiproliferative; antiviral; ss.

XX Mus sp.

XX WO955728-A2.

XX 04-NOV-1999.

PD

XX PF 27-APR-1999; 99WO-CA000375.
 XX XX 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX (HSCR-) HSC RES & DEV LP.
 PA Egan SE, Wang W, Sengar A;
 PI WPI: 2000-052802/04.
 DR P-PSDB; AAY57449.
 XX New nucleic acid encoding Esei and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 PS Claim 6; Page 56-59; 99pp; English.
 CC The present invention specifically describes mammalian Esei and 2
 CC proteins (1) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (1) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Esp1s protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (1) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (1)-specific antibodies (Ab);
 CC sequences antisense to the (1) polynucleotide; agents that downregulate
 CC expression of Ese genes or antagonists of an Ese binding partner are used
 CC to treat diseases associated with undesirable overexpression of Esei and
 CC changes in cellular function. Particularly overexpression of Esei is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (1) is used to promote endocytosis of selected cells.
 CC (Ant)agonists of (1) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Ese mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Ese-Esp1s complex, then binding dynamin to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signaling, tissue development or synaptic transmission. The
 CC present sequence represents mouse EseiL cDNA sequence
 CC
 SQ Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.02e-34 Length: 5738
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_740_800 (1-61) x AA239024 (1-5738)
 QY 1 VallysValValTyTYrArGAlaLeuTYrProPhgIuSeArGserHisAspGluile 20
 DB 2471 GTAAAGAGGTATATTACCGAGCGCTGTACCCCTTGATTCAGAAATCAGATGAGATC 2530
 QY 21 Thr1legInProGlyAsp1leValMetValAspGluSeGlnThnGlyGluuproglyTTP 40
 DB 2531 ACCATCCAGCCAGGAGATTAATGTCATGTGGATGAAGCAAGCTGGAAGCCAGGATGG 2590
 QY 41 LeuGlyGlyGluLeuLeuGlyHisThrGlyTTPPhProAlaAsnTyZalaglulysile 60
 DB 2591 CTTGAGAGGAGACTGAAAGGAGAGACGGATGGTTCCTCGCAAACTATGCAAGAAAGATT 2650
 QY 61 Pro 61
 DB 2651 CCA 2653
 RESULT 12
 AA234571
 ID AA234571 standard; cDNA; 5458 BP.
 XX AA234571;
 AC

XX DT 01-FEB-2000 (first entry)
 XX XX Human SH3D1A cDNA clone 21.
 XX DE SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 XX KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW hematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 267..3929
 FT CDS /**tag= a
 FT
 XX WO9953062-A2.
 XX 21-OCT-1999.
 XX 16-APR-1999; 99WO-US008371.
 XX 16-APR-1998; 98US-0082007P.
 XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 PA Korenberg JR, Chen X;
 PI WPI: 1999-633829/54.
 DR P-PSDB; AAY32155.
 XX Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 PS Claim 2; Fig 6; 99pp; English.
 XX This is the nucleotide sequence of full-length cDNA (clone 21)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukaemia, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21 (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC association of gains in chromosome 21 with leukaemia, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 CC
 SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.81e-32 Length: 5458
 Score: 320.50 Matches: 61
 Percent Similarity: 92.428 Conservative: 0
 Best Local Similarity: 92.428 Mismatches: 0
 Query Match: 96.25% Indels: 5
 DB: Gaps: 1
 US-09-720-934-2_COPY_740_800 (1-61) x AA234571 (1-5458)

QY 1 ValIysValIYrTYrArgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
 Db 2484 GTAAAGTGTGTATTACCGGCACTGTACCCCTTTGATCCAGAAAGCATGATGAATC 2543
 QY 21 ThrIleGlnProGlyAspIleValMet-----ValAspGluSerGlnThr 35
 Db 2344 ACTATCCAGCCAGAGACATAGTCATGTTAAAGGGGAATGGTGGATGAAGCCAAACT 2603
 QY 36 GlyGluProGlyTYrPLeuGlyGlyGluLeuYsGlyLysThrGlyTYrPheProAlaAsn 55
 Db 2604 GGAAACCCGCGCTGCTTGAGAGAAATTAAAGAAAGACAGGCTGTTCCCTGCAGAAC 2663
 QY 56 TyrAlaGluYsIlePro 61
 Db 2664 TATGCAGAGAAATCCCA 2681

RESULT 13
 AAS84763
 ID AAS84763 standard; cDNA; 7435 BP.
 XX
 AC AAS84763;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #20567.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 PN MO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US008631.
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 XX
 P-PSDB; ABG20576.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID NO 20567; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polynucleotide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.66e-32 Length: 7435
 Score: 320.50 Matches: 61
 Percent Similarity: 92.42% Conservative: 0
 Best Local Similarity: 92.42% Mismatches: 0
 Query Match: 96.25% Indels: 5
 DB: Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x AAS84763 (1-7435)

QY 1 ValIysValIYrTYrArgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
 Db 2499 GTAAAGTGTGTATTACCGGCACTGTACCCCTTTGATCCAGAAAGCATGATGAATC 2558
 QY 21 ThrIleGlnProGlyAspIleValMet-----ValAspGluSerGlnThr 35
 Db 2559 ACTATCCAGCCAGAGACATAGTCATGTTAAAGGGGAATGGTGGATGAAGCCAAACT 2618
 QY 36 GlyGluProGlyTYrPLeuGlyGlyGluLeuYsGlyLysThrGlyTYrPheProAlaAsn 55
 Db 2619 GGAAACCCGCGCTGCTTGAGAGAAATTAAAGAAAGACAGGCTGTTCCCTGCAGAAC 2678
 QY 56 TyrAlaGluYsIlePro 61
 Db 2679 TATGCAGAGAAATCCCA 2696

RESULT 14
 AAT39799
 ID AAT39799 standard; DNA; 2873 BP.
 XX
 AC AAT39799;
 XX
 DT 19-FEB-1998 (first entry)
 XX
 DE Human clone 65 gene.
 XX
 KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
 KW cellular signalling element; cellular structural element; malignancy;
 KW protein identification; functional domain; protein screening;
 KW cellular signal transduction process; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..1532
 FT /*tag= a
 XX
 PN MO9631625-A1.
 PD 10-OCT-1996.
 XX
 PF 04-APR-1996; 96WO-US004454.
 PR 07-APR-1995; 95US-00417872.
 PR 03-APR-1996; 96US-00630915.
 XX
 PA (CYTO-) CYTOGEN CORP.
 PA (UYN-) UNITV NORTH CAROLINA.
 PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;
 XX
 DR WPI: 1996-465045/46.
 XX
 P-PSDB; AAW05399.

Identifying polypeptide(s) having specific functional domain (esp. SH3 domain) - comprises detecting selective binding to recognition unit, regardless of sequence homology.

PS Claim 76; Fig 58; 174pp; English.
CC AAT39786-T39803 represent novel human and mouse genes encoding Src-
CC homology region 3 (SH3) domain containing proteins that can be used in
CC the method of the invention. SH3 domain containing proteins play a role
CC in signalling and structural elements of cells. The method of the
CC invention is for identifying polypeptides containing functional domains
CC of interest (especially SH3 domains). The method comprises contacting a
CC multivalent recognition unit (RU) complex with a number of peptides and
CC identifying polypeptides having a selective binding affinity for the RU
CC complex. The method is based on functional similarities and does not rely
CC on sequence similarities. Prior methods only gave limited success for
CC identifying proteins which contain an SH3 domain due to the minimal
CC sequence homology among known SH3 proteins. It has been found that small
CC peptide RUs in multivalent form have reduced specificity for a given
CC functional domain compared to monomer RUs. Multivalent RU complexes are
CC particularly suited to screening for polypeptides containing functional
CC domains that are similar to, but not identical in sequence to, the
CC original target functional domain. The new method enables proteins having
CC a common function to be identified. Identification of novel SH3 proteins
CC will be useful for a better understanding of cell growth, malignancy,
CC signal transduction processes, etc. New candidate drugs can be
CC identified, and their specificities (e.g. pharmacological activities) can
CC be assessed using the method of the invention
CC
CC
SQ Sequence 2873 BP; 903 A; 506 C; 623 G; 841 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.09e-19 Length: 2873
Score: 216.00 Matches: 36
Percent Similarity: 74.58% Conservative: 8
Best Local Similarity: 61.02% Mismatches: 15
Query Match: 64.86% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AAT39799 (1-2873)
QY 3 ValValIYTYTYRAGAlALeuTYrProBheGluSerArgSerHisASpGluIleThrIle 22
Db 228 TTGGTGAATTATAGACATTAACCCCTTTGAAAGCAACCATGATGAGTGGTTT 287
QY 23 GlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYrPleugly 42
Db 288 AATTCTGAGATATATATTCAGGTTCATGAAAAAACCGAGAGAACCTGGTGGCTTAT 347
QY 43 GlyGluLeuGlyGlyeThrGlyTYrPheProAlaAsnTYrAlaGluIlePro 61
Db 348 GGTAGTTTTCAGGAAATTTTGCTGGTTCATGCAATATATGTAGAAAAATGCCA 404

RESULT 15
AAZ39011
ID AAZ39011 standard; cDNA; 3593 BP.
AC AAZ39011;
XX
XX 28-FEB-2000 (first entry)
DT
XX
XX Mouse E2e2 coding sequence.
DE
XX
XX Mouse; murine; E2e1; E2e2; endocytosis; vesicular trafficking;
KM regulation; actin cytoskeleton; detection; cancer; infection;
KM EH-domain and SH3-domain regulator of endocytosis; anticancer;
KM antiproliferative; antiviral; 88.
XX
XX
XX Mus sp.
OS
XX
XX W09955728-A2.
PN
XX
XX 04-NOV-1999.
PD
XX
XX 27-APR-1999; 99WO-CA000375.
PF
XX
XX 27-APR-1998; 96CA-02230201.
PR

PR 05-FEB-1999; 99US-0118739P.
XX
XX (HSCR-) HSC RES & DEV LP.
PA
XX
XX Egan SE, Wang W, Sengar A;
PI
XX
XX WPI; 2000-052802/04.
DR
XX
XX P-PSDB; AAY57445.
PT
XX
XX New nucleic acid encoding E2e1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
PS
XX
XX Claim 25; Page 46-48; 99pp; English.
CC
CC The present sequence encodes mouse E2e2. The present invention
CC specifically describes mammalian E2e1 and 2 proteins (I) and their splice
CC variants (E2e = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with E2e1 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of E2e genes or
CC antagonists of an E2e binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of E2e1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) is used to promote endocytosis of selected cells. (ant)agonists of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive E2e mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an E2e-E2e1 complex, then binding dynamin to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
CC
CC
SQ Sequence 3593 BP; 1171 A; 783 C; 850 G; 789 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.2e-18 Length: 3593
Score: 216.00 Matches: 36
Percent Similarity: 75.86% Conservative: 8
Best Local Similarity: 62.07% Mismatches: 14
Query Match: 64.86% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AAZ39011 (1-3593)
QY 3 ValValIYTYTYRAGAlALeuTYrProBheGluSerArgSerHisASpGluIleThrIle 22
Db 2155 TTGGTGAATTATAGACATTAACCCCTTTGAAAGCAACCATGATGAGTGGTTT 2214
QY 23 GlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYrPleugly 42
Db 2215 AGTTCGTGAGATATATATTCAGGTTCATGAAAAAACCTGAGAGAACCTGGTGGCTTAT 2274
QY 43 GlyGluLeuGlyGlyeThrGlyTYrPheProAlaAsnTYrAlaGluIlePro 60
Db 2275 GGTAGTTTTCAGGAAATTTTGCTGGTTCATGCAATATATGTAGAAAAATGTG 2328

Search completed: July 1, 2004, 13:31:47
Job time : 199.286 secs

(UYNC-) UNIV NORTH CAROLINA.
 PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;
 DR WPI; 1996-465045/46.
 DR P-PSDB; AAM05395.
 XX
 PT Identifying polypeptide(s) having specific functional domain (esp. SH3
 PT domain) - comprises detecting selective binding to recognition unit,
 PT regardless of sequence homology.
 XX
 PS Claim 76; Fig 50; 174pp; English.
 XX
 CC AAT39786-T39803 represent novel human and mouse genes encoding Src-
 CC homology region 3 (SH3) domain containing proteins that can be used in
 CC the method of the invention. SH3 domain containing proteins play a role
 CC in signalling and structural elements of cells. The method of the
 CC invention is for identifying polypeptides containing functional domains
 CC of interest (especially SH3 domains). The method comprises contacting a
 CC multivalent recognition unit (RU) complex with a number of peptides and
 CC identifying polypeptides having a selective binding affinity for the RU
 CC complex. The method is based on functional similarities and does not rely
 CC on sequence similarities. Prior methods only gave limited success for
 CC identifying proteins which contain an SH3 domain due to the minimal
 CC sequence homology among known SH3 proteins. It has been found that small
 CC peptide RUs in multivalent form have reduced specificity for a given
 CC functional domain compared to monomer RUs. Multivalent RU complexes are
 CC particularly suited to screening for polypeptides containing functional
 CC domains that are similar to, but not identical in sequence to, the
 CC original target functional domain. The new method enables proteins having
 CC a common function to be identified. Identification of novel SH3 proteins
 CC will be useful for a better understanding of cell growth, malignancy,
 CC signal transduction processes, etc. New candidate drugs can be
 CC identified, and their specificities (e.g. pharmacological activities) can
 CC be assessed using the method of the invention
 CC
 XX SQ Sequence 1389 BP; 404 A; 370 C; 357 G; 258 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,01e-36 Length: 1389
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_1080_1138 (1-59) x AAT39795 (1-1389)
 QY 1 AAlaValCyGlnValIleGlyNetyrAspYrTrhAlaGlnAsnAspAspGluLeuAla 20
 Db 1192 GGAGGTGTGCGAGGATGTTGGATGTGACGCTACCGCCGAGATGACATGAGCTGGCC 1251
 QY 21 PheAsnLySGlyGlnIleLeuValLeuAsnLySGluAspProAspTrpTrpLySGly 40
 Db 1252 TTCACACAGGCGCCAGTATCATCAACGTCCTCAACAGAGAGACCCCTGACGTGGAAAGGA 1311
 QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTrpVallyLeuThrThrAsp 59
 Db 1312 GAAGTCATAGGACAGAGTGGGCTCTTCCATTCATCATTTATGTGAAGTGCACAGAC 1368
 RESULT 2
 AB055007
 ID AB055007 standard; cDNA; 2067 BP.
 AC AB055007;
 XX
 XX 22-AUG-2002 (first entry)
 DE Human ovarian antigen HVBVX28 cDNA, SEQ ID NO:887.
 XX Human ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 21q22.1-22.2;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN W020020677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001WO-US018569.
 XX
 PR 07-JUN-2000; 2000US-0209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147876/19.
 DR P-PSDB; ABP41930.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 PS Claim 1; SEQ ID NO 887; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP3228) and to cDNAs encoding them (AB054131-AB056305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2067 BP; 614 A; 426 C; 418 G; 605 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: 3.32e-36 Length: 2067
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

```

US-09-720-934-2_COPY_1080_1138 (1-59) x ABQ55007 (1-2067)
QY 1 A1aValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
Db 299 GCAAGTGTCCAGGCGATGGATGGATGACACCGCGAGATGAGATGAGATGGCC 358
QY 21 PheAsnlysglyGlnIleIleAsnValLeuAsnlysglyAspProAspTyrTrpIlysgly 40
Db 359 TTCAACAAGAGGCGCATCATCAGCTCTCAACAAGAGGAGCCTGACTGCTGGAAGA 418
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValIlyLeuThrThrAsp 59
Db 419 GAAGTCATATGACAAAGTGGGCTCTCCATCCATCAATTATGTGAAGCTGACACAGAC 475
RESULT 3
AAZ34574
ID AAZ34574 standard; cDNA; 3231 BP.
XX
AC AAZ34574;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA clone 9.
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 2..1927
FT FT /*tag= a
XX
EN WO953062-A2.
XX
PD 21-OCT-1999.
XX
PF 16-APR-1999; 99WO-US008371.
XX
PR 16-APR-1998; 98US-0082007P.
XX
PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
PI Korenberg JR, Chen X;
XX
DR WPI; 1999-633829/54.
DR P-PSDB; AAY32158.
XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
PT diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
PS Claim 2; Fig 14; 99pp; English.
XX
XX
XX This is the nucleotide sequence of a non-full-length cDNA (clone 9)
XX corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
XX contributes to the development of platelets and the pathogenesis of
XX leukaemias, both in general and in particular those involving the
XX megakaryocytic lineage. The SH3D1A gene maps to the small candidate
XX region for low platelets on chromosome 21. Sequencing of 5 different
XX sizes of cDNA clone from foetal brain (see AAZ34570-74) suggests that at
XX least 3 isoforms exist. The invention provides methods for the diagnosis
XX and treatment of megakaryocytic abnormality, myeloproliferative disorder,
XX platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
XX association of gains in chromosome 21, low platelets in deletion for 21,
XX abnormalities, dysfunctions and disorders including brain malformations
XX and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
XX and colpocephaly. Methods are also provided for: suppressing cells unable
XX to regulate themselves; screening for a somatic alteration in the SH3D1A
XX gene; monitoring the progress and adequacy of a treatment; monitoring

```

```

CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
CC treatment of a subject (including a prenatal subject) having
CC megakaryocytic abnormality, myeloproliferative disorder, platelet
CC disorder, leukaemia or neural disorder using a nucleic acid that
CC expresses SH3D1A or its antisense nucleic acid
XX
SQ Sequence 3231 BP; 1004 A; 721 C; 712 G; 794 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 5,95e-16 Length: 3231
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
XX
US-09-720-934-2_COPY_1080_1138 (1-59) x AAZ34574 (1-3231)
QY 1 A1aValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
Db 1730 GCAAGTGTCCAGGCGATGGATGGATGACACCGCGAGATGAGATGAGATGGCC 1789
QY 21 PheAsnlysglyGlnIleIleAsnValLeuAsnlysglyAspProAspTyrTrpIlysgly 40
Db 1790 TTCAACAAGAGGCGCATCATCAGCTCTCAACAAGAGGAGCCTGACTGCTGGAAGA 1849
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValIlyLeuThrThrAsp 59
Db 1850 GAAGTCATATGACAAAGTGGGCTCTCCATCCATCAATTATGTGAAGCTGACACAGAC 1906
RESULT 4
AAZ34572
ID AAZ34572 standard; cDNA; 5195 BP.
XX
AC AAZ34572;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA clone 11.
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 239..3886
FT FT /*tag= a
XX
EN WO953062-A2.
XX
PD 21-OCT-1999.
XX
PF 16-APR-1999; 99WO-US008371.
XX
PR 16-APR-1998; 98US-0082007P.
XX
PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
PI Korenberg JR, Chen X;
XX
DR WPI; 1999-633829/54.
DR P-PSDB; AAY32156.
XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
PT diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
PS Claim 2; Fig 10; 99pp; English.
XX
XX

```

CC This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA34570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis,
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukaemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid

XX Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;

SQ

Alignment Scores:

Pred. No.:	1.07e-35	Length:	5195
Score:	323.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x AA234572 (1-5195)

QY 1 AAlaValCyGGInValIIleGlyMeCTyrAspYrThrAlaGlnAsnAspAspGluLeuAla 20

Db 3689 GCAGGTGCCAGGTATGGAGTGTACGACTACACCGCCAGCAATGACGATGAGCTGGCC 3748

QY 21 PheAsnLysGlyGlnIleleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40

Db 3749 TTCAACAAGGGCCGATCATCAACGTCTCCAAACAAGAGAGACCTGACTGGTGAAGAAGA 3808

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrVallyleuThrThrAsp 59

Db 3809 GAAGTCATGACAGAGTGGGGCTCTCCATCCCAATTATGTGAAGCTGACCAAGAC 3865

RESULT 5

AAZ34570

ID AA234570 standard; cDNA; 5199 BP.

AC AAZ34570;

DT 01-FEB-2000 (first entry)

XX Human SH3D1A cDNA.

DE

XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;

XX megakaryocytic abnormality; myeloproliferative disorder;

KW platelet disorder; neural disorder; thrombocytopenia;

KM haematopoietic disorder; cognitive dysfunction; microcephaly;

KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.

XX Homo sapiens.

OS

XX Key Location/Qualifiers

FH CDS 208..3642

FT /tag= a

XX WO9953062-A2.

PN 21-OCT-1999.

XX PD

XX 16-APR-1999; 99WO-US008371.

PF 16-APR-1998; 98US-0082007P.

PR (CEDA-) CEDARS SINAI HEALTH SYSTEM.

PA Korenberg JR, Chen X;

XX WPI; 1999-633829/54.

DR P-PSDB; AAY32154.

XX Nucleic acid from the human SH3D1A gene and its products, useful for the

PT diagnosis and treatment of myeloproliferative disorders and leukemia.

PS Claim 2; Fig 5; 99Pp; English.

XX This is the nucleotide sequence of full-length cDNA corresponding to a

CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the

CC development of platelets and the pathogenesis of leukaemias, both in

CC general and in particular those involving the megakaryocytic lineage. The

CC SH3D1A gene maps to the small candidate region for low platelets on

CC chromosome 21. Sequencing of 5 different sizes of cDNA clone (see

CC AA234570-74) suggests that at least 3 isoforms exist. The invention

CC provides methods for the diagnosis and treatment of megakaryocytic

CC abnormality, myeloproliferative disorder, platelet disorder, acute

CC leukaemia, neural disorders, thrombocytopenia, platelet disorder on

CC chromosome 21, low platelets in deletion for 21, association of gains in

CC chromosome 21 with leukaemias, neural abnormalities, dysfunctions and

CC disorders including brain malformations and corresponding cognitive

CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are

CC also provided for: suppressing cells unable to regulate themselves;

CC screening for a somatic alteration in the SH3D1A gene; monitoring the

CC progress and adequacy of a treatment; monitoring tumour risk progress or

CC megakaryocytic abnormality, myeloproliferative disorder, haematopoietic

CC disorder, platelet disorder or leukaemia; and treatment of a subject

CC (including a prenatal subject) having megakaryocytic abnormality,

CC myeloproliferative disorder, platelet disorder, leukaemia or neural

CC disorder using a nucleic acid that expresses SH3D1A or its antisense

CC nucleic acid

XX Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

SQ

Alignment Scores:

Pred. No.:	1.07e-35	Length:	5199
Score:	323.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x AA234570 (1-5199)

QY 1 AAlaValCyGGInValIIleGlyMeCTyrAspYrThrAlaGlnAsnAspAspGluLeuAla 20

Db 3445 GCAGGTGCCAGGTATGGAGTGTACGACTACACCGCCAGCAATGACGATGAGCTGGCC 3504

QY 21 PheAsnLysGlyGlnIleleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40

Db 3505 TTCAACAAGGGCCGATCATCAACGTCTCCAAACAAGAGAGACCTGACTGGTGAAGAAGA 3564

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrVallyleuThrThrAsp 59

Db 3565 GAAGTCATGACAGAGTGGGGCTCTCCATCCCAATTATGTGAAGCTGACCAAGAC 3621

RESULT 6

AAZ34571

ID AAZ34571 standard; cDNA; 5458 BP.

XX AAZ34571;

AC

XX 01-FEB-2000 (first entry)

DT

XX

DE Human SH3D1A cDNA clone 21.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 OS Homo sapiens.
 XX
 XX
 FH Key location/Qualifiers
 FT CDS 267..3929
 FT /*tag= a
 XX
 XX WO9953062-A2.
 XX
 XX 21-OCT-1999.
 XX
 XX 16-APR-1999; 99WC-US008371.
 XX
 XX 16-APR-1998; 98US-0082007P.
 XX
 XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 XX Korenberg JR, Chen X;
 XX
 XX WPI; 1999-633829/54.
 XX P-PSDB; AAY32155.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 PS
 PS Claim 2; Fig 8; 99pp; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 21)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see A4234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC association of gains in chromosome 21, low platelets in deletion for 21,
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.14e-35 Length: 5458
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-720-934-2_COPY_1080_1138 (1-59) x AA234571 (1-5458)
 QY 1 A1AVal1CysG1Val1IleG1MetTyrAspTyrTrpAlaGlnAsnAspAspGluLeuAla 20
 Db 3732 GCAAGTGTCCAGAGGATTTGGATGTACGACTACCGCGCGAAGTACGAGTACGCTGCGC 3791

QY 21 PheAsnIysG1Gln1Ile1LeuAsnValLeuAsnIysG1uAspProAspTyrTrpLysGly 40
 Db 3792 TTCAACAAGGCGCCAGATCATCAAGCTCTCACAAAGAGAGACCTGATGTTGAAAGGA 3851
 QY 41 GluValAsnG1Gln1Val1GlyLeuPheProSerAsnTyrVal1LysLeuThrThrasp 59
 Db 3852 GAAGTCAATGACAAAGTGGGCTCTTCCATCCATATATGTGAAGCTACCCACAGAC 3908
 RESULT 7
 AAS84762
 ID AAS84762 standard; cDNA; 2874 BP.
 XX
 XX AAS84762;
 XX
 XX 13-FEB-2002 (first entry)
 XX
 XX DNA encoding novel human diagnostic protein #20566.
 DE
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 XX W0200175067-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 XX
 XX 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI; 2001-639362/73.
 XX P-PSDB; ABG20575.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 XX Claim 1; SEQ ID NO 20566; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2874 BP; 940 A; 679 C; 703 G; 552 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.9e-35 Length: 2874

Score: 319.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 98.76%
DB: 5

Matches: 58
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AAS84762 (1-2874)

QY 2 ValAlaGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21

DB 1855 GTGTGCGAGGTGATGTGATGTAAGACTACACCGCAGATGACGATGAGCTGGCTTC 1914

QY 22 AsnLysGlyGlnIleIleAsnValIleAsnLysGlnAspProAspTyrPlyGlyGlu 41

DB 1915 AACAGGCGCAGATCAACAGCTCTCAACAGAGGACCTGCTGCTGGAAGAGAA 1974

QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrAsp 59

DB 1975 GTCAATGACAGTGGGCTCTTCCCATCCATTATGTGAAGCTGACACAGAC 2028

RESULT 8
AAZ39009 standard; cDNA; 3723 BP.

XX AC AAZ39009;

XX DT 28-FEB-2000 (first entry)

XX DE Mouse Esel coding sequence.

XX KM Mouse; murine; Esel; Eseq2; endocytosis; vesicular trafficking;

XX KM EH-domain; actin cytoskeleton; detection; cancer; infection;

XX KM antiproliferative; antiviral; ss.

XX OS Mus sp.

XX PN WO9955728-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-CA000375.

XX PR 27-APR-1998; 98CA-02230201.

XX PR 05-FEB-1999; 99US-0118739P.

XX PA (HSCR-) HSC RES & DEV LP.

XX PI Egan SE, Wang W, Sengar A;

XX DR WPI; 2000-052802/04.

XX DR P-PSDB; AAY57444.

XX PS New nucleic acid encoding Esel and 2 proteins, involved in regulation of

XX PT endocytosis, used e.g. for treating cancer or preventing viral infection.

XX XX Claim 6; Page 40-42; 99pp; English.

XX The present sequence encodes mouse Esel. The present invention

XX specifically describes mammalian Esel and 2 proteins (1) and their splice

XX variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)

XX are involved in regulation of clathrin-mediated endocytosis (as a complex

XX with Eps15 protein), vesicular trafficking and actin cytoskeleton.

XX Generally (1) or its (ant)agonists, mimetics, fragments and inactive

XX mutants; (1)-specific antibodies (Ab); sequences antisense to the (1)

XX polynucleotide; agents that downregulate expression of Ese genes or

XX antagonists of an Ese binding partner are used to treat diseases

XX associated with undesirable endocytosis and resulting changes in cellular

XX function. Particularly overexpression of Esel is used to block clathrin-

XX mediated endocytosis in vivo or in cell cultures, while administration of

XX (1) is used to promote endocytosis of selected cells. (Ant)agonists of

XX (1) or Ab are used to suppress abnormal proliferation of cells that can

XX be stimulated to proliferate by a growth factor receptor; and similar

CC compounds (also inactive Ese mutants) can be used to prevent viral

CC infection. Endocytosis may also be regulated, in vivo or in cell

CC cultures, by forming an Ese-Eps15 complex, then binding dynam to the

CC complex. Generally conditions that can be treated include cancer;

CC abnormal cell division or migration; viral infection; or abnormal

CC receptor signalling, tissue development or synaptic transmission

QY Alignment Scores: 1.92e-34 length: 3723

QY 1 AlAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20

DB 3445 GCAGTGTGCCAGGTGATCGGATGTAAGATTACACCGCCAGAACGATGACGAATGACC 3504

QY 21 PheAsnLysGlyGlnIleIleAsnValIleAsnLysGlnAspProAspTyrPlyGlyGlu 40

DB 3505 TTCAGCAAGGCCAGATCATCAACGCTCTCAACAGAGGAGACCGGACTGTGGAAGAA 3564

QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrAsp 59

DB 3565 GAAGTCAGTGGGACAGTGGGCTCTTCCCATCCATTATGTGAAGCTGACACAGAC 3621

RESULT 9
AAZ39008 standard; cDNA; 5082 BP.

XX AC AAZ39008;

XX DT 28-FEB-2000 (first entry)

XX DE Mouse Esel full length cDNA sequence.

XX KM Mouse; murine; Esel; Eseq2; endocytosis; vesicular trafficking;

XX KM EH-domain; actin cytoskeleton; detection; cancer; infection;

XX KM antiproliferative; antiviral; ss.

XX OS Mus sp.

XX PN WO9955728-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-CA000375.

XX PR 27-APR-1998; 98CA-02230201.

XX PR 05-FEB-1999; 99US-0118739P.

XX PA (HSCR-) HSC RES & DEV LP.

XX PI Egan SE, Wang W, Sengar A;

XX DR WPI; 2000-052802/04.

XX DR P-PSDB; AAY57444.

XX PS New nucleic acid encoding Esel and 2 proteins, involved in regulation of

XX PT endocytosis, used e.g. for treating cancer or preventing viral infection.

XX XX Claim 6; Page 38-40; 99pp; English.

XX The present sequence encodes mouse Esel. The present invention

XX specifically describes mammalian Esel and 2 proteins (1) and their splice

XX variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)

XX are involved in regulation of clathrin-mediated endocytosis (as a complex

CC with BspI5 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) or Ab are used to promote endocytosis of selected cells. (ant)agonists of
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-BspI5 complex, then binding dynamn to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX

SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2, 84e-34 Length: 5082
Score: 313.00 Matches: 57
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.61% Mismatches: 0
Query Match: 96.90% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AA239008 (1-5082)

QY 1 AAlaValCyGlnValIleGlyMetTyRaspTyRThAlaGlnAsnAspGluLeuAla 20
DB 3702 GGAAGTGTCCAGGTGATCGGATGTACGATTACCGCCAGAACATGACGACTAGCC 3761
QY 21 PheAsnIysGlyGlnIleIleAsnValIleuAsnIysGlyAspProAspTyrTrpIysGly 40
DB 3762 TTCAGCAAGGCGCAGATCATCAACGCTCCACCAAGAGAGACCGGACTGTGGAAAGGA 3821
QY 41 GluValaAsnGlyGlnValaGlyLeuPheProSerAsnTyRValIysLeuThrThrasp 59
DB 3822 GAAATCAGTGGCGCAAGTTGGGCTCTTCCATCCAAATTATGTAAAGCTGACCAAGAC 3878

RESULT 10
AA239025
ID AA239025 standard; cDNA; 5144 BP.
XX
AC AA239025;
XX
DT 28-FEB-2000 (first entry)
XX

DE Mouse Ese1l coding sequence.
XX
XX Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
XX MO9955728-A2.
XX
XX 04-NOV-1999.
XX
XX 27-APR-1999; 99WO-CA000375.
XX
XX 27-APR-1998; 98CA-02230201.
XX
XX 05-FEB-1999; 99US-0118739P.
XX
XX (HSCR-) HSC RES & DEV LP.
XX
XX Egan SE, Wang W, Sengar A;
XX
XX WPI; 2000-052802/04.

DR P-PSDB: AA575449.
XX
XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
PT
XX
XX
PS Claim 6; Page 55-62; 99pp; English.

CC The present invention specifically describes mammalian Ese1 and 2
CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
CC regulator of endocytosis). (I) are involved in regulation of clathrin-
CC mediated endocytosis (as a complex with BspI5 protein), vesicular
CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
CC sequences antisense to the (I) polynucleotide; agents that downregulate
CC expression of Ese genes or antagonists of an Ese binding partner are used
CC to treat diseases associated with undesirable endocytosis and resulting
CC changes in cellular function. Particularly overexpression of Ese1 is used
CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
CC administration of (I) or Ab are used to promote endocytosis of selected cells.
CC (ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
CC cells that can be stimulated to proliferate by a growth factor receptor;
CC and similar compounds (also inactive Ese mutants) can be used to prevent
CC viral infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-BspI5 complex, then binding dynamn to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission. The
XX present sequence represents mouse Ese1l coding sequence

SQ Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2, 89e-34 Length: 5144
Score: 313.00 Matches: 57
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.61% Mismatches: 0
Query Match: 96.90% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AA239025 (1-5144)

QY 1 AAlaValCyGlnValIleGlyMetTyRaspTyRThAlaGlnAsnAspGluLeuAla 20
DB 3445 GCAAGTGTCCAGGTGATCGGATGTACGATTACCGCCAGAACATGACGACTAGCC 3504
QY 21 PheAsnIysGlyGlnIleIleAsnValIleuAsnIysGlyAspProAspTyrTrpIysGly 40
DB 3505 TTCAGCAAGGCGCAGATCATCAACGCTCTCCATCCAAATTATGTAAAGCTGACCAAGAC 3564

QY 41 GluValaAsnGlyGlnValaGlyLeuPheProSerAsnTyRValIysLeuThrThrasp 59
DB 3565 GAAATCAGTGGCGCAAGTTGGGCTCTTCCATCCAAATTATGTAAAGCTGACCAAGAC 3621

RESULT 11
AA239024
ID AA239024 standard; cDNA; 5738 BP.
XX
AC AA239024;
XX
XX 28-FEB-2000 (first entry)
XX
XX
XX Mouse Ese1l cDNA sequence.
XX
XX
XX Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX
XX
XX Mus sp.
XX
XX MO9955728-A2.
XX
XX 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.
XX PF
XX 27-APR-1998; 98CA-02230201.
XX PR
XX 05-FEB-1999; 99US-0118739P.
XX PR
XX (HSCR-) HSC RES & DEV LP.
XX PA
XX Egan SE, Wang W, Sengar A;
XX PI
XX WPI; 2000-052802/04.
XX DR
XX P-PSDB; AAY57449.
XX PT
XX New nucleic acid encoding Ees1 and 2 proteins, involved in regulation of
XX endocytosis, used e.g. for treating cancer or preventing viral infection.
XX PT
XX
XX PS
XX Claim 6; Page 56-59; 99p; English.
XX
XX The present invention specifically describes mammalian Ees1 and 2
XX proteins (I) and their splice variants (Ees = EH-domain and SH3-domain
XX regulator of endocytosis). (I) are involved in regulation of clathrin-
XX mediated endocytosis (as a complex with Ees1 protein), vesicular
XX trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
XX mimetics, fragments and inactive mutants); (II)-specific antibodies (Ab);
XX sequences antisense to the (I) polynucleotide; agents that downregulate
XX expression of Ees genes or antagonists of an Ees binding partner are used
XX to treat diseases associated with undesirable overexpression of Ees1 is used
XX changes in cellular function. Particularly overexpression of Ees1 is used
XX to block clathrin-mediated endocytosis in vivo or in cell cultures, while
XX administration of (I) is used to promote endocytosis of selected cells.
XX (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
XX cells that can be stimulated to proliferate by a growth factor receptor;
XX and similar compounds (also inactive Ees mutants) can be used to prevent
XX viral infection. Endocytosis may also be regulated, in vivo or in cell
XX cultures, by forming an Ees-Ees1 complex, then binding dynamin to the
XX complex. Generally conditions that can be treated include cancer;
XX abnormal cell division or migration that can be treated include cancer;
XX receptor signaling, tissue development or synaptic transmission. The
XX present sequence represents mouse Ees1L cDNA sequence
XX
XX SQ
XX Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3.32e-34 Length: 5738
XX Score: 313.00 Matches: 57
XX Percent Similarity: 100.00% Conservative: 2
XX Best Local Similarity: 96.61% Mismatches: 0
XX Query Match: 96.90% Indels: 0
XX DB: Gaps: 0
XX
XX US-09-720-934-2_COPY_1080_1138 (1-59) x AA239024 (1-5738)
XX
XX QY 1 A1aValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
XX DB 3704 GCAGGTGCGCAGTGTATCGGATGTAACGATACACCGCCCAAGATACCAACTAGCC 3763
XX QY 21 PheAsnLysGlyGlnIleLeuValLeuAsnLysGluAspProAspTyrTrpLysGly 40
XX DB 3764 TTCACCAAGGCGCAGATCATCAACGCTCTCAACAGAGAGACCCGACGCTGGGAAAG 3823
XX QY 41 GluValaAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
XX DB 3824 GAAGTCAAGGCGCAGTGTGAGTCTTCCATCCATTAATGTAAGCTGACCAAGAC 3880
XX
XX RESULT 12
XX ID AAS84763 standard; cDNA; 7435 BP.
XX AC AAS84763;
XX XX
XX DT 13-FEB-2002 (first entry)
XX XX DNA encoding novel human diagnostic protein #20567.

XX XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX KW
XX Homo sapiens.
XX OS
XX WO200175067-A2.
XX PN
XX 11-OCT-2001.
XX PD
XX 30-MAR-2001; 2001WO-US008631.
XX PF
XX 31-MAR-2000; 2000US-00540217.
XX PR
XX 23-AUG-2000; 2000US-00649167.
XX XX
XX (HYSE-) HYSEQ INC.
XX PA
XX Dimaac RT, Liu C, Tang YT;
XX PT
XX WPI; 2001-639362/73.
XX DR
XX P-PSDB; ABG20576.
XX DR
XX
XX PT
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensic, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX PT
XX
XX PS
XX Claim 1; SEQ ID NO 20567; 103p; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridization probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful for treating disorders
XX of sites expressing (II). (I) and (II) are useful for identifying disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensic, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ
XX Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.24e-33 Length: 7435
XX Score: 310.00 Matches: 59
XX Percent Similarity: 98.33% Conservative: 0
XX Best Local Similarity: 98.33% Mismatches: 0
XX Query Match: 95.98% Indels: 1
XX DB: Gaps: 0
XX
XX US-09-720-934-2_COPY_1080_1138 (1-59) x AAS84763 (1-7435)
XX
XX QY 1 A1aValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
XX DB 3758 GCAGGTGCGCAGTGTATCGGATGTAACGATACACCGCCCAAGATACCAACTAGCC 3817
XX QY 21 PheAsnLys-GlyGlnIleLeuValLeuAsnLysGluAspProAspTyrTrpLysGly 40
XX DB 3818 TTCACCAAGGCGCAGATCATCAACGCTCTCAACAGAGAGACCCGACGCTGGGAAAG 3877
XX QY 40 YGluValaAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
XX

DB 3878 AGAAGTCATGACAGTGGGCTCTTCCATCCATTATGTGAAGCTGACCAAGAC 3935

RESULT 13

AAT39796

ID AAT39796 standard; DNA; 747 BP.

XX

AC AAT39796;

XX

DT 19-FEB-1998 (first entry)

XX

DE Human SH3P18 gene.

XX

KM Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;

KM cellular signalling element; cellular structural element; malignancy;

KM protein identification; functional domain; protein screening;

KM cellular signal transduction process; ss.

XX

OS Homo sapiens.

XX

FH Key

FT Location/Qualifiers

FT 1..747

FT CDS /*tag= a

FT /product= "SH3P18"

XX

PN MO9631625-A1.

XX

PD 10-OCT-1996.

XX

PF 04-APR-1996; 96MO-US004454.

XX

PR 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

XX

PA (CYTO-) CYTOGEN CORP.

PA (UNYC-) UNIV NORTH CAROLINA.

XX

PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX

DR WPI; 1996-465045/46.

DR P-PSDB; AAW05396.

XX

PT Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

XX

PS Claim 76; Fig 52; 174pp; English.

XX

CC AAT39786-139803 represent novel human and mouse genes encoding Src-

CC homology region 3 (SH3) domain containing proteins that can be used in

CC the method of the invention. SH3 domain containing proteins play a role

CC in signalling and structural elements of cells. The method of the

CC invention is for identifying polypeptides containing functional domains

CC of interest (especially SH3 domains). The method comprises contacting a

CC multivalent recognition unit (RU) complex with a number of peptides and

CC identifying polypeptides having a selective binding affinity for the RU

CC complex. The method is based on functional similarities and does not rely

CC on sequence similarities. Prior methods only gave limited success for

CC identifying proteins which contain an SH3 domain due to the minimal

CC sequence homology among known SH3 proteins. It has been found that small

CC peptide RUs in multivalent form have reduced specificity for a given

CC functional domain compared to monomer RUs. Multivalent RU complexes are

CC particularly suited to screening for polypeptides containing functional

CC domains that are similar to, but not identical in sequence to, the

CC original target functional domain. The new method enables proteins having

CC a common function to be identified. Identification of novel SH3 proteins

CC will be useful for a better understanding of cell growth, malignancy,

CC signal transduction processes, etc. New candidate drugs can be

CC identified, and their specificities (e.g. pharmacological activities) can

CC be assessed using the method of the invention

XX

SQ Sequence 747 BP; 244 A; 133 C; 186 G; 184 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.41e-28 Length: 747

Score: 266.00 Matches: 44

Percent Similarity: 91.38% Conservative: 9

Best Local Similarity: 75.86% Mismatches: 5

Query Match: 82.35% Indels: 0

DB: 2 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AAT39796 (1-747)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAlaPhe 21

DB 553 GTATGTGAGGTGATTTGCTATGTATGACTATGACGCAAAATATGAAATGAGCTCAGTTTC 612

QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrPylsGlyGlu 41

DB 613 TCCAGGACACACTCATTTATGTATGACAAAGATATCTGATTGGTGGCAAGAGAG 672

QY 42 ValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrAsp 59

DB 673 ATCAAGGGGTGACTGGTCTCTTCCCTCAAACTACGTTAAGATGACGACAGAC 726

RESULT 14

ID AAT39799 standard; DNA; 2873 BP.

XX

AC AAT39799;

XX

DT 19-FEB-1998 (first entry)

XX

DE Human clone 65 gene.

XX

KM Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;

KM cellular signalling element; cellular structural element; malignancy;

KM protein identification; functional domain; protein screening;

KM cellular signal transduction process; ss.

XX

OS Homo sapiens.

XX

FH Key

FT Location/Qualifiers

FT 3..1532

FT CDS /*tag= a

XX

PN MO9631625-A1.

XX

PD 10-OCT-1996.

XX

PF 04-APR-1996; 96MO-US004454.

XX

PR 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

XX

PA (CYTO-) CYTOGEN CORP.

PA (UNYC-) UNIV NORTH CAROLINA.

XX

PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX

DR WPI; 1996-465045/46.

DR P-PSDB; AAW05399.

XX

PT Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

XX

PS Claim 76; Fig 58; 174pp; English.

XX

CC AAT39786-139803 represent novel human and mouse genes encoding Src-

CC homology region 3 (SH3) domain containing proteins that can be used in

CC the method of the invention. SH3 domain containing proteins play a role

CC in signalling and structural elements of cells. The method of the

CC invention is for identifying polypeptides containing functional domains

CC of interest (especially SH3 domains). The method comprises contacting a

CC multivalent recognition unit (RU) complex with a number of peptides and

CC identifying polypeptides having a selective binding affinity for the RU

CC complex. The method is based on functional similarities and does not rely
CC on sequence similarities. Prior methods only gave limited success for
CC identifying proteins which contain an SH3 domain due to the minimal
CC sequence homology among known SH3 proteins. It has been found that small
CC peptide fusions in multivalent form have reduced specificity for a given
CC functional domain compared to monomer fusions. Multivalent RU complexes are
CC particularly suited to screening for polypeptides containing functional
CC domains that are similar to, but not identical in sequence to, the
CC original target functional domain. The new method enables proteins having
CC a common function to be identified. Identification of novel SH3 proteins
CC will be useful for a better understanding of cell growth, malignancy,
CC signal transduction processes, etc. New candidate drugs can be
CC identified, and their specificities (e.g. pharmacological activities) can
CC be assessed using the method of the invention

US-09-720-934-2_COPY_1080_1138 (1-59) x AAT39799 (1-2873)

Alignment Scores:

Pred. No.:	7.79e-28	Length:	2873
Score:	266.00	Matches:	44
Percent Similarity:	91.38%	Conservative:	9
Best Local Similarity:	75.86%	Mismatches:	5
Query Match:	82.35%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x AAT39799 (1-2873)

QY 2 ValCysGlnValIleGlyMetLysPyrThrAlaGlnAsnAspArgGlnLeuAlaPhe 21

Db 1338 GTATGTCAGGATGATGCTATGATGACATGACACAAATATGATGAAGATGACCTCAGTTTC 1397

QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGlnAspProAspTrpTrpLysGlyGlu 41

Db 1398 TCCAAAGGACACATCATTAATGATGAACAAGATGATCTGATGATGGCAAGAGAG 1457

QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTrpValLysLeuThrThrAsp 59

Db 1458 ATCAAAGGGGTGACTGCTCTCTTCTTCTTCAACATGATGAAGTACACAGACAGAC 1511

RESULT 15

ID ABK43586 standard; cDNA, 3746 BP.

XX ABK43586;

AC

XX 05-JUN-2002 (first entry)

DT

XX

DE DNA encoding novel central nervous system protein #166.

XX

KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasia; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminization;
KW endocrine disorder; diabetes; cancer; leukemia; neovascularization;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

XX WO200155318-A2.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001332.

XX

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214866P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0233422P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234297P.

PR 25-SEP-2000; 2000US-0234988P.

PR 25-SEP-2000; 2000US-0234989P.

PR 26-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 27-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0253678P.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX MPI; 2001-581633/65.
XX P-PSDB; AAU87256.
PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives or
XX preservatives.
PS Claim 1; SEQ ID NO 176; 837bp; English.
XX The invention describes an isolated nucleic acid molecule (1) encoding a

CC novel central nervous system protein. (1) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angioneuroma, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphasia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Alignment Scores:

Pred. No.:	1.09e-27	Length:	3746
Score:	266.00	Matches:	44
Percent Similarity:	91.38%	Conservative:	9
Best Local Similarity:	75.86%	Mismatches:	5
Query Match:	82.35%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x ABR43586 (1-3746)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAspAspGluLeuAlaPhe 21
Db 1264 GATATGACGGTATGTTGCTATGTTATGCTATGACGCAAAATATATGAAGTACGCTGAGTTTC 1323
QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGlu 41
Db 1324 TCCAAAGGACAACTCATATGTTATGTAAGCAAAAGATGATCTATTTGGTGCAGAGAGG 1383
QY 42 ValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThraSp 59
Db 1384 ATCAACGGGGTGACTGTCTCTTCTTCAAACTAGCTTAAGATGACAGACAGC 1437

Search completed: July 1, 2004, 13:32:11
Job time : 189.112 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 11:41:07 ; Search time 197.546 Seconds

(without alignments)
1376.313 Million cell updates/sec

Title: US-09-720-934-2_COPY_999_1062
341

Sequence: 1 GEEIAGVIAATATGATGPEQLT.....RGKKRQIGWPEANVVKLSP 64

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.P/US09720934/funat_30062004_064539_13422/app_query.fasta_1.1386
-DB=N.Geneseq.236rand4 -QWRT=fastlap -SUFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITs=b1ts -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934 @CGN 1.1 1596 @runat_30062004_064539_13422 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq.29Jan04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	341	100.0	1389	2	AAT39795 Human SH3
2	341	100.0	3231	2	AAZ34574 Human SH3
3	341	100.0	5199	2	AAZ34570 Human SH3
4	330	96.8	5195	2	AAZ34572 Human SH3
5	330	96.8	5458	2	AAZ34571 Human SH3
6	325	95.3	3723	3	AAZ39009 Mouse Ese
7	325	95.3	5082	3	AAZ39008 Mouse Ese
8	325	95.3	5144	3	AAZ39025 Mouse Ese

9	325	95.3	5738	3	AAZ39024 Mouse Ese
10	319	93.5	2067	6	ABQ55007 Human cva
11	303	88.9	7435	5	AAS84763 DNA encod
12	253	74.2	747	2	AAT39796 Human SH3
13	253	74.2	2873	2	AAT39799 Human SH3
14	253	74.2	3593	3	AAZ39011 Human SH3
15	253	74.2	3746	4	ABK43586 DNA encod
16	253	74.2	4210	9	AAS27090 CDNA encod
17	253	74.2	4210	9	ADB93268 Human CDN
18	253	74.2	4447	4	AAS02055 Human CDN
19	253	74.2	4625	3	AAZ39010 Human SH3
20	253	74.2	4975	6	AAZ39027 Human SH3
21	253	74.2	5828	6	AAZ39027 Human SH3
22	253	74.2	6014	3	AAZ39026 Human SH3
23	253	74.2	6103	4	AAZ39026 Human SH3
24	237	69.5	3981	4	ABL01995 Human SH3
25	237	69.5	7225	4	ABL01994 Human SH3
26	202	59.2	2874	5	AAS84762 DNA encod
27	163	47.8	395	5	AAS84759 DNA encod
28	118.5	34.8	3111	4	ABL02239 Drosophil
29	118.5	34.8	8061	4	ABL02238 Drosophil
30	116.5	34.2	2971	9	ADD71209 Human int
31	116.5	34.2	3204	5	ABA09727 Human int
32	116.5	34.2	4804	6	ABK3624 Human bon
33	113.5	33.3	423	3	AAC01364 Human CDN
34	113.5	33.3	720	3	AAA44672 Human sec
35	113.5	33.3	1215	4	AAH41499 Human sec
36	113.5	33.3	1254	4	AAH41499 Human sec
37	113.5	33.3	1254	4	AAH41499 Human sec
38	113.5	33.3	1392	2	AAT79834 Human tyr
39	113.5	33.3	1392	2	AAT79834 Human tyr
40	113.5	33.3	1392	2	AAT79834 Human tyr
41	113.5	33.3	1392	2	AAT79834 Human tyr
42	113.5	33.3	1392	2	AAT79834 Human tyr
43	113.5	33.3	1659	2	AAT79834 Human tyr
44	113.5	33.3	1659	2	AAT79834 Human tyr
45	113.5	33.3	1659	2	AAT79834 Human tyr

ALIGNMENTS

RESULT 1	AAT39795 standard; DNA; 1389 BP.
AC	AAT39795;
DT	19-FEB-1998 (first entry)
DE	Human SH3P17 gene.
XX	Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
KW	cellular signalling element; cellular structural element; malignancy;
KW	protein identification; functional domain; protein screening;
KW	cellular signal transduction process; ss.
OS	Homo sapiens.
XX	
FT	Key
FT	Location/Qualifiers
FT	1..1389
FT	/tag=a
FT	/product="SH3P17"
XX	
XX	WO9631625-A1.
XX	
XX	10-OCT-1996.
XX	
XX	04-APR-1996; 96WO-US004454.
XX	
XX	07-APR-1995; 95US-00417872.
XX	
XX	03-APR-1996; 96US-00630915.
XX	
XX	(CYTO-) CYTOGEN CORP.


```

QY 21 LeuAlaProGlyGlnLeuIleLeuIleArgLysAsnProGlyGlyTTPTrpGlnGly 40
Db 1547 CTCGCCCTGCTCAGCTGATTTGATCCGAAAAAGAACCCAGTGGATGTTGGAGGA 1606
QY 41 GlnLeuGlnAlaArgGlyLysLysArgGlnIleGlyTTPheProAlaAsnTyValLys 60
Db 1607 GAGCTGCAGAGCAGCGGGGAAAAAGCCAGATAGGCTGTTCCCGCTAAATTATGTAAG 1666
QY 61 LeuLeuSerPro 64
Db 1667 CTTCTAAGCCCT 1678
RESULT 3
AAZ34570
ID AAZ34570 standard; cDNA; 5199 BP.
XX
AC AAZ34570;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA.
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 208..3642
FT /*tag= a
XX
PN WO9953062-A2.
XX
PD 21-OCT-1999.
XX
PF 16-APR-1999; 99WO-US008371.
XX
PR 16-APR-1998; 98US-0082007P.
XX
PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
PI Korenberg JR, Chen X;
XX
DR WPI: 1999-633829/54.
XX
DR P-PSDB; AAY32154.
XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
PT diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
PS Claim 2; Fig 5; 99pp; English.
XX
CC This is the nucleotide sequence of full-length cDNA corresponding to a
CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the
CC development of platelets and the pathogenesis of leukemias, both in
CC general and in particular those involving the megakaryocytic lineage. The
CC SH3D1A gene maps to the small candidate region for low platelets on
CC chromosome 21. Sequencing of 5 different sizes of cDNA clone (see
CC AAZ34570-74) suggests that at least 3 isoforms exist. The invention
CC provides methods for the diagnosis and treatment of megakaryocytic
CC abnormality, myeloproliferative disorder, platelet disorder, acute
CC leukaemia, neural disorders, thrombocytopenia, platelet disorder on
CC chromosome 21, low platelets in deletion for 21, association of gains in
CC disorders including brain malformations and corresponding cognitive
CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
CC also provided for: suppressing cells unable to regulate themselves;
CC screening for a somatic alteration in the SH3D1A gene; monitoring the
CC progress and adequacy of a treatment; monitoring tumour risk progress or
CC megakaryocytic abnormality; myeloproliferative disorder; haematopoietic
CC disorder, platelet disorder or leukaemia; and treatment of a subject

```

```

CC (including a prenatal subject) having megakaryocytic abnormality,
CC myeloproliferative disorder, platelet disorder, leukaemia or neural
CC disorder using a nucleic acid that expresses SH3D1A or its antisense
CC nucleic acid
SQ Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2-Re-32 Length: 5199
Score: 341.00 Matches: 64
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-09-720-934-2_copy_999_1062 (1-64) x AAZ34570 (1-5199)
QY 1 GlnGlnGlnIleAlaGlnValIleAlaSerTyrThrAlaThrGlyProGlnGlnLeuThr 20
Db 3202 GGAGAGAGAAATTGCCAGTTATTCCTCATACACCGCCACCGGCCCGAGCACTCACT 3261
QY 21 LeuAlaProGlyGlnLeuIleLeuIleArgLysLysAsnProGlyGlyTTPTrpGlnGly 40
Db 3262 CTCGCCCTGCTCAGCTGATTTGATCCGAAAAAGAACCCAGTGGATGTTGGAGGA 3321
QY 41 GlnLeuGlnAlaArgGlyLysLysArgGlnIleGlyTTPheProAlaAsnTyValLys 60
Db 3322 GAGCTGCAGAGCAGTGGGAAAAAGCCAGATAGGCTGTTCCCGCTAAATTATGTAAG 3381
QY 61 LeuLeuSerPro 64
Db 3382 CTTCTAAGCCCT 3393
RESULT 4
AAZ34572
ID AAZ34572 standard; cDNA; 5195 BP.
XX
AC AAZ34572;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA clone 11.
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 239..3886
FT /*tag= a
XX
PN WO9953062-A2.
XX
PD 21-OCT-1999.
XX
PF 16-APR-1999; 99WO-US008371.
XX
PR 16-APR-1998; 98US-0082007P.
XX
PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
PI Korenberg JR, Chen X;
XX
DR WPI: 1999-633829/54.
XX
DR P-PSDB; AAY32156.
XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
PT diagnosis and treatment of myeloproliferative disorders and leukemia.
XX

```

PS Claim 2; Fig 10; 99pp; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukemias, neural
 CC abnormalities, dysfunctional dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 CC
 XX
 SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6.69e-31 Length: 5195
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 Gaps: 0
 DB: 2
 US-09-720-934-2_COPY_999_1062 (1-64) x AA234572 (1-5195)
 QY 3 GtlltlealaglnvalillealaserTyrrhralathrglyProgluglnleuthleuAla 22
 Db 3452 GAATTCGCCAGGTTATTCCTCATACACCGCACCGGCCCGAGACACTCTCCGC 3511
 QY 23 Proglglnleuilleuileuileuileuileuileuileuileuileuileuileu 42
 Db 3512 CCGGTGACGCTGATTTGATCGGAAAAAGAACCCAGGTGGATGGAGAGAGCTG 3571
 QY 43 GlnlaaarglylylsyAsnProglgylTrrpPheproAlaAsnTyValylsleu 62
 Db 3572 CAAGCAGCGGGAAAAAGCGCGAGTAGGCTGTTCCACGCTAATTATGTAAGCTTCTA 3631
 QY 63 SerPro 64
 Db 3632 AGCCT 3637
 RESULT 5
 AA234571
 ID AA234571 standard; CDNA; 5458 BP.
 XX
 AC AA234571;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 21.
 XX
 KM SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KM megakaryocytic abnormality; myeloproliferative disorder;
 KM platelet disorder; neural disorder; thrombocytopenia;
 KM haematopoietic disorder; cognitive dysfunction; microcephaly;
 KM lissencephaly; colpocephaly; helioproensephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers

FT CDS 267..3929
 FT /*tag= a
 XX
 PN MO9953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99MO-US008371.
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 DR WPI; 1999-633829/54.
 DR P-PSDB; AAY32155.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 8; 99pp; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 21)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukemias, neural
 CC abnormalities, dysfunctional dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 CC
 XX
 SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 7.12e-31 Length: 5458
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 Gaps: 0
 DB: 2
 US-09-720-934-2_COPY_999_1062 (1-64) x AA234571 (1-5458)
 QY 3 GtlltlealaglnvalillealaserTyrrhralathrglyProgluglnleuthleuAla 22
 Db 3495 GAATTCGCCAGGTTATTCCTCATACACCGCACCGGCCCGAGACACTCTCCGC 3554
 QY 23 Proglglnleuilleuileuileuileuileuileuileuileuileuileuileu 42
 Db 3555 CCGGTGACGCTGATTTGATCGGAAAAAGAACCCAGGTGGATGGAGAGAGCTG 3614
 QY 43 GlnlaaarglylylsyAsnProglgylTrrpPheproAlaAsnTyValylsleu 62
 Db 3615 CAAGCAGCGGGAAAAAGCGCGAGTAGGCTGTTCCACGCTAATTATGTAAGCTTCTA 3674
 QY 63 SerPro 64
 DB: 2

DB 3675 AGCCT 3680

RESULT 6
AAZ39009
ID AAZ39009 standard; cDNA; 3723 BP.

AC AAZ39009;
XX
XX
XX 28-FEB-2000 (first entry)
XX
DE Mouse Ese1 coding sequence.
XX
XX Mouse; murine; Ese1, Ese2; endocytosis; vesicular trafficking;
XX regulation; actin cytoskeleton; detection; cancer; infection;
XX EH-domain and SH3-domain regulator of endocytosis; anticancer;
XX antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
XX WO9955728-A2.
XX
XX 04-NOV-1999.
XX
XX 27-APR-1999; 99WO-CA000375.
XX
XX 27-APR-1998; 98CA-02230201.
XX 05-FEB-1999; 99US-0118739P.
XX
XX (HSCR-) HSC RES & DEV LP.
XX
XX Egan SE, Wang W, Sengar A;
XX
XX WPI; 2000-052802/04.
XX P-PSDB; AAY57444.
XX
XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
XX endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
XX
XX Claim 6; Page 40-42; 99pp; English.

The present sequence encodes mouse Ese1. The present invention
specifically describes mammalian Ese1 and 2 proteins (I) and their splice
variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
are involved in regulation of clathrin-mediated endocytosis (as a complex
with Esp15 protein), vesicular trafficking and actin cytoskeleton.
Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
polynucleotide; agents that downregulate expression of Ese genes or
antagonists of an Ese binding partner are used to treat diseases
associated with undesirable endocytosis and resulting changes in cellular
function. Particularly overexpression of Ese1 is used to block clathrin-
mediated endocytosis in vivo or in cell cultures, while administration of
(I) is used to promote endocytosis of selected cells. (Ant)agonists of
(I) or Ab are used to suppress abnormal proliferation of cells that can
be stimulated to proliferate by a growth factor receptor; and similar
compounds (also inactive Ese mutants) can be used to prevent viral
infection. Endocytosis may also be regulated, in vivo or in cell
cultures, by forming an Ese-Esp15 complex, then binding dynamitin to the
complex. Generally conditions that can be treated include cancer;
abnormal cell division or migration; viral infection; or abnormal
receptor signalling, tissue development or synaptic transmission

XX
XX Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,87e-30 Length: 3723
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AAZ39009 (1-3723)

QY 3 Glu1leaIagInVal11leaISerTyThrAlaThrglyProgluGlnThreua1a 22
DB 3208 GAAATTGCCAGTATTATGCTCTCAGCTGCTACTGCTCCGAAACAACCTGAGCT 3267
QY 23 ProgluGlnleu1leaIeArglyLysaAnProgluGlyTTPTrpGluG1eGluDen 42
DB 3268 CCGGAGAGCTGATTCGATCCGAAAAAGAACCCAGGTGATGATGGAGAGAACTG 3327
QY 43 Glu1laArglyLysaArgGln1eG1yThrpheProAlaAntTyRVal1ysleuDen 62
DB 3328 CAGCTCGAGGGAAGAAAGCCAGATAGGGTGTTCCAGCAAAATATGTCMAACTTCTA 3387
QY 63 SerPro 64
DB 3388 AGCCCT 3393

RESULT 7
AAZ39008
ID AAZ39008 standard; cDNA; 5082 BP.
XX
XX AAZ39008;
XX
XX 28-FEB-2000 (first entry)
XX
XX Mouse Ese1 full length cDNA sequence.
XX
XX Mouse; murine; Ese1, Ese2; endocytosis; vesicular trafficking;
XX regulation; actin cytoskeleton; detection; cancer; infection;
XX EH-domain and SH3-domain regulator of endocytosis; anticancer;
XX antiproliferative; antiviral; ss.
XX
XX Mus sp.
XX
XX WO9955728-A2.
XX
XX 04-NOV-1999.
XX
XX 27-APR-1999; 99WO-CA000375.
XX
XX 27-APR-1998; 98CA-02230201.
XX 05-FEB-1999; 99US-0118739P.
XX
XX (HSCR-) HSC RES & DEV LP.
XX
XX Egan SE, Wang W, Sengar A;
XX
XX WPI; 2000-052802/04.
XX P-PSDB; AAY57444.
XX
XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
XX endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
XX
XX Claim 6; Page 38-40; 99pp; English.

The present sequence encodes mouse Ese1. The present invention
specifically describes mammalian Ese1 and 2 proteins (I) and their splice
variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
are involved in regulation of clathrin-mediated endocytosis (as a complex
with Esp15 protein), vesicular trafficking and actin cytoskeleton.
Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
polynucleotide; agents that downregulate expression of Ese genes or
antagonists of an Ese binding partner are used to treat diseases
associated with undesirable endocytosis and resulting changes in cellular
function. Particularly overexpression of Ese1 is used to block clathrin-
mediated endocytosis in vivo or in cell cultures, while administration of
(I) is used to promote endocytosis of selected cells. (Ant)agonists of
(I) or Ab are used to suppress abnormal proliferation of cells that can
be stimulated to proliferate by a growth factor receptor; and similar
compounds (also inactive Ese mutants) can be used to prevent viral
infection. Endocytosis may also be regulated, in vivo or in cell
cultures, by forming an Ese-Esp15 complex, then binding dynamitin to the

CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2.76e-30 Length: 5082
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x AAZ39008 (1-5082)
QY 3 Glu1le1aGlnVal1le1a1aSerTyrThra1aThrGlyProGlnGlnLeu1a1a 22
Db 3465 GAATATGCCAGGATTGCTTCTACGCTGCTACTGCTCCGAAACAACCTCAGCTG 3524
QY 23 ProGlyGlnLeu1leu1a1aGlyLysAsnProGlyGlyTyrTPGlnGlyGlnLeu 42
Db 3525 CCGGGCAGCTGATTGCTGATCCGAAAAAGAACCCAGTGGATGGAGGAAGACTG 3584
QY 43 Gln1a1aArgGlyLysLysArgGln1leGlyTyrPhePro1aAsnTyrValLysLeu1e 62
Db 3585 CAAGCTCGAGGGAAAAAGCGCCAGATAGGGTGTTCACGCAAAATATGTCAACTCTTA 3644
QY 63 SerPro 64
Db 3645 AGCCCC 3650
RESULT 8
AAZ39025
ID AAZ39025 standard; cDNA; 5144 BP.
XX
AC AAZ39025;
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse Esei1 coding sequence.
XX
KW Mouse; murine; Esei1; Esei2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
PN WO9955728-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
DR WPI; 2000-052802/04.
DR P-PSDB; AAY57449.
XX
PT New nucleic acid encoding Esei1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS Claim 6; Page 59-62; 9pp; English.
XX
CC The present invention specifically describes mammalian Esei1 and 2
CC proteins (I) and their splice variants (See = EH-domain and SH3-domain
CC regulator of endocytosis). (I) are involved in regulation of clathrin-

CC mediated endocytosis (as a complex with Eps15 protein), vesicular
CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
CC mimetics, antagonists to the (I) polynucleotide; agents that downregulate
CC expression of Esei genes or antagonists of an Esei binding partner are used
CC to treat diseases associated with undesirable overexpression of Esei1 used
CC changes in cellular function. Particularly overexpression of Esei1 used
CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
CC administration of (I) is used to promote endocytosis of selected cells.
CC (ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
CC cells that can be stimulated to proliferate by a growth factor receptor;
CC and similar compounds (also inactive Esei mutants) can be used to prevent
CC viral infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Esei-Eps15 complex, then binding dynamin to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission. The
CC present sequence represents mouse Esei1 coding sequence
XX
SQ Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2.8e-30 Length: 5144
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x AAZ39025 (1-5144)
QY 3 Glu1le1aGlnVal1le1a1aSerTyrThra1aThrGlyProGlnGlnLeu1a1a 22
Db 3208 GAATATGCCAGGATTGCTTCTACGCTGCTACTGCTCCGAAACAACCTCAGCTG 3267
QY 23 ProGlyGlnLeu1leu1a1aGlyLysAsnProGlyGlyTyrTPGlnGlyGlnLeu 42
Db 3268 CCGGGCAGCTGATTGCTGATCCGAAAAAGAACCCAGTGGATGGAGGAAGACTG 3327
QY 43 Gln1a1aArgGlyLysLysArgGln1leGlyTyrPhePro1aAsnTyrValLysLeu1e 62
Db 3328 CAAGCTCGAGGGAAAAAGCGCCAGATAGGGTGTTCACGCAAAATATGTCAACTCTTA 3387
QY 63 SerPro 64
Db 3388 AGCCCC 3393
RESULT 9
AAZ39024
ID AAZ39024 standard; cDNA; 5738 BP.
XX
AC AAZ39024;
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse Esei1 cDNA sequence.
XX
KW Mouse; murine; Esei1; Esei2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
PN WO9955728-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX

PA (HSCR-) HSC RES & DEV LP.
 XX
 XX Egan SE, Wang W, Sengar A;
 XX
 XX WPI: 2000-052802/04.
 DR P-PSDB: AAY57449.
 XX
 XX
 PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 PS Claim 6; Page 56-59; 99pp; English.
 XX
 CC The present invention specifically describes mammalian Esei and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Eps15 protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC sequences antisense to the (I) polynucleotide; agents that downregulate
 CC expression of Esei genes or antagonists of an Esei binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Esei is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Esei mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Esei-Eps15 complex, then binding dynam to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The
 CC present sequence represents mouse Esei cDNA sequence
 XX
 SQ Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3, 21e-30 Length: 5738
 Score: 325.00 Matches: 61
 Percent Similarity: 98.39% Conservative: 0
 Best Local Similarity: 98.39% Mismatches: 1
 Query Match: 95.31% Indels: 0
 Gaps: 0
 DB: 3
 US-09-720-934-2_COPY_999_1062 (1-64) x AA239024 (1-5738)
 QY 3 GUILLEALGIVALLLEALSERLYRTHALATHGLYPRGIGLNLNTHLEUALA 22
 DB 3467 GAATATGCCAGGTATGCTTCTTACGCTACTGATGCTCCGAAACACTCACTGGCT 3526
 QY 23 PROGLGNLNLNLEULEULEALGLYSLYASNPGLYGLYTRPRTPGLNUGLYGULEN 42
 DB 3527 CCGGGACACTGATTTGATCCGAAAAGAACCCGAGGAGATGGTGGAGAGAACTG 3586
 QY 43 GINALAARGLYLYSLYSLYARGINILEGLYTRPPEPRALAAANTYVALLYSLEULEN 62
 DB 3587 CAAGCTCGAGGAAAAGAGCCAGATAGAGGTGTTCCAGCAATTATCTCAACTTCTA 3646
 QY 63 Serpo 64
 DB 3647 AGCCCC 3652
 RESULT 10
 ABQ55007
 ID ABQ55007 standard; cDNA; 2067 BP.
 XX
 AC ABQ55007;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 XX Human ovarian antigen HVBX28 cDNA, SEQ ID NO:887.
 DE Human ovarian antigen, ovary; ovarian; breast; cancer; tumour;
 XX
 KW Human; ovarian antigen, ovary; ovarian; breast; cancer; tumour;

KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antineoplastic; gynaecological; reproductive; chromosome 21q22.1-22.2;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200200677-A1.
 XX
 XX 03-JAN-2002.
 XX
 XX 07-JUN-2001; 2001WO-US018569.
 XX
 XX 07-JUN-2000; 2000US-0209467P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Birse CE, Rosen CA;
 XX
 XX WPI: 2002-147878/19.
 DR P-PSDB: ABP41930.
 DR
 XX
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 XX
 PS Claim 1; SEQ ID NO 887; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2067 BP; 614 A; 426 C; 418 G; 605 T; 0 U; 4 Other;
 XX
 Alignment Scores:
 Pred. No.: 5, 07e-30 Length: 2067
 Score: 319.00 Matches: 59
 Percent Similarity: 98.39% Conservative: 2
 Best Local Similarity: 95.16% Mismatches: 1
 Query Match: 93.55% Indels: 0

DB:	6	Gaps:	0
US-09-720-934-2_COPY_999_1062 (1-64)	x	AB055007 (1-2067)	
Qy	3	GluIleAlaGlnValIleAlaSerYrThrIaThnGlyProGluGlnLeuThrIleuAla	22
	62	AAATTGCCCCAGGTTATTTCCTCATACACCGCCACCGGCCGAGAGCTCACTCGGC	121
Db	23	ProGlyGlnLeuIleLeuIleArgIysIysAsnProGlyGlyTrpTrpGluGlyIleu	42
	122	CCTGTCACGCTGATTTTGTATCCGAAAAAGAACCCAGCGTGATGGTGGGAGAGAGCTG	181
Qy	43	GlnAlaIaGgGlyLysIysArgGlnIleGlyTrpPheProAlaAsnYrValIysIleu	62
Db	182	CAAGCACGGGAGAAAAAGCCAGATAGGCTGGTCCAGCTAATATGTAAGCTTCTA	241
Qy	63	SerPro 64	
	242	AGCCCT 247	
Db			
RESULT 11			
AA884763			
ID	AA884763	standard; cDNA; 7435 BP.	
XX	AA884763;		
XX	13-FEB-2002	(first entry)	
DT			
XX			
XX	DNA encoding novel human diagnostic protein #20567.		
DE			
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder; se.		
OS	Homo sapiens.		
XX			
XX	WO200175067-A2.		
PN			
PD	11-OCT-2001.		
XX			
PF	30-MAR-2001; 2001WO-US008631.		
XX			
PR	31-MAR-2000; 2000US-00540217.		
PR	23-AUG-2000; 2000US-00649167.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Dmanac RT, Liu C, Tang YT;		
XX			
DR	WPI; 2001-639362/73.		
XX	P-Psdb; ABG20576.		
XX			
PT	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics, forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity.		
XX			
XX	Claim 1; SEQ ID NO 20567; 103bp; English.		
XX			
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)		
CC	sequences. (I) is useful as hybridisation probes, polymerase chain		
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping.		
CC	and in recombinant production of (II). The polynucleotides are also used		
CC	in diagnostics as expressed sequence tags for identifying expressed		
CC	genes. (I) is useful in gene therapy techniques to restore normal		
CC	activity of (II) or to treat disease states involving (II). (II) is		
CC	useful for generating antibodies against it, detecting or quantitating a		
CC	polypeptide in tissue, as molecular weight markers and as a food		
CC	supplement. (II) and its binding partners are useful in medical imaging		
CC	of sites expressing (II). (I) and (II) are useful for treating disorders		
CC	involving aberrant protein expression or biological activity. The		
CC	polypeptide and polynucleotide sequences have applications in		
CC	diagnostics, forensics, gene mapping, identification of mutations		
CC	responsible for genetic disorders or other traits to assess biodiversity		

```
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA64197-AAA94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPD at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 7435 BF; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      2.54e-27          Length:      7435
Score:          303.00           Matches:     58
Percent Similarity: 96.83%       Conservative: 3
Best Local Similarity: 92.06%    Mismatches:  1
Query Match:     88.86%         Indels:     1
DB:              5               Gaps:       0

US-09-720-934-2_COPY_999_1062 (1-64) x AAS84763 (1-7435)

QY      3  GHUlealeAGInValllleaIasEryThrAlathrgGlyProGIugInLeuthrleuA1a 22
Db      3520 AAAATTGCCCGCAGGTATTTGGCTCTCAACACC GCCCGCCGCGACGCTACTCTGC CC 3579
QY      23  ProGIugInLeutleuleleArGLySAsn-ProGIgLyTPTrpGIugLyJule 42
Db      3580 CCTGGTCAGCTGATTTTGATTCGA AAAAGA ACCCGCAGGTGATGTGGAAAGAGAGCT 3639
QY      42  uGHAlAArgGLyLySLysArGgInllEGLYTYPhePrOAlAantYrValLysleule 62
Db      3640 GCAAGCACAGTGGGAAAAAGCCGACGATGTGGCTGTTCACAGCTAATTAATGAAGCTTCT 3699
QY      62  uSerPro 64
Db      3700 AACCCCT 3706

RESULT 12
ID      AAT39796 standard; DNA; 747 BP.
AC      AAT39796;
DT      19-FEB-1998 (first entry)
DE      Human SH3P18 gene.
KW      Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
        cellular signaling element; cellular structural element; malignancy;
        protein identification; functional domain; protein screening;
        cellular signal transduction process; ss.
OS      Homo sapiens.
FH      Key Location/Qualifiers
FT      CDS 1..747
FT      /tag=a
FT      /product= "SH3P18"
XX      WO9631625-A1.
XX      PD 10-OCT-1996.
XX      PF 04-APR-1996; 96MO-US004454.
XX      PR 07-APR-1995; 95US-00417872.
XX      PR 03-APR-1996; 96US-00630915.
XX      PA (CYTO-) CYTOGEN CORP.
XX      PA (UYNC-) UNIV NORTH CAROLINA.
XX      Sparks AB, Hofman N, Kay BK, Fowlkes DM, McConnell SJ,
XX      PFI 1996-465045/46.
DR      P-PADB; AAM05396.
```

XX Identifying polypeptide(s) having specific functional domain (esp. SH3 domain) - comprises detecting selective binding to recognition unit, regardless of sequence homology.

PS Claim 76; Fig 52; 174pp; English.

CC AAT39786-T39803 represent novel human and mouse genes encoding Src-homology region 3 (SH3) domain containing proteins that can be used in the method of the invention. SH3 domain containing proteins play a role in signalling and structural elements of cells. The method of the invention is for identifying polypeptides containing functional domains of interest (especially SH3 domains). The method comprises contacting a multivalent recognition unit (RU) complex with a number of peptides and identifying polypeptides having a selective binding affinity for the RU complex. The method is based on functional similarities and does not rely on sequence similarities. Prior methods only gave limited success for identifying proteins which contain an SH3 domain due to the minimal sequence homology among known SH3 proteins. It has been found that small functional domain compared to monomer RUs. Multivalent RU complexes are particularly suited to screening for polypeptides containing functional domains that are similar to, but not identical in sequence to, the original target functional domain. The new method enables proteins having a common function to be identified. Identification of novel SH3 proteins will be useful for a better understanding of cell growth, malignancy, signal transduction processes, etc. New candidate drugs can be identified, and their specificities (e.g. pharmacological activities) can be assessed using the method of the invention.

SO Sequence 747 BP; 244 A; 133 C; 186 G; 184 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	747
Score:	253.00	47
Percent Similarity:	87.10%	7
Best Local Similarity:	75.81%	8
Query Match:	74.19%	Indels: 0
DB:	2	Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AAT39796 (1-747)

QY 3 GlnlelaaglnvalilleaSerTyThralathrglyProgluglnleuthleuAla 22

Db 334 GAGATTGCTCAGGTAACCTCAGCATATGTTGCTTCTGCTGACAACTTACCTTGCA 393

QY 23 Progluglnleuileuileuileuileuileuileuileuileuileuileuileu 42

Db 394 CAGGACAGTTAATTAATTTCTAAAGAAAATACAGGTGGTGGCAAGAGTTA 453

QY 43 GlnlaaArgGlyLysLysArgGlnleuileuileuileuileuileuileuileu 62

Db 454 CAGGCAAGAGAAAAGGCAAGAAAGGATGTTCTGCGCAATGTTAACTTTTG 513

QY 63 SerPro 64

Db 514 GGTCCA 519

RESULT 13

AAT39799

ID AAT39799 standard; DNA; 2873 BP.

XX AAT39799;

AC AAT39799;

XX 19-FEB-1998 (first entry)

XX Human clone 65 gene.

XX Src-homology region 3 domain; human; mouse; SH3 domain; cell growth; cellular signalling element; cellular structural element; malignancy; protein identification; functional domain; protein screening; cellular signal transduction process; ss.

OS Homo sapiens.

XX Key

XX Location/Qualifiers

FT 3..1532

FT CDS

XX /**tag= a

XX W09631625-A1.

XX 10-0CT-1996.

XX 04-APR-1996; 96MO-US004454.

XX 07-APR-1995; 95US-00417872.

XX 03-APR-1996; 96US-00630915.

XX (CYTO-) CYTOGEN CORP.

XX (YUNC-) UNIV NORTH CAROLINA.

PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX WPI; 1996-465045/46.

DR P-PSDB; AAM05399.

XX

PT Identifying polypeptide(s) having specific functional domain (esp. SH3 domain) - comprises detecting selective binding to recognition unit, regardless of sequence homology.

XX Claim 76; Fig 58; 174pp; English.

CC AAT39786-T39803 represent novel human and mouse genes encoding Src-homology region 3 (SH3) domain containing proteins that can be used in the method of the invention. SH3 domain containing proteins play a role in signalling and structural elements of cells. The method of the invention is for identifying polypeptides containing functional domains of interest (especially SH3 domains). The method comprises contacting a multivalent recognition unit (RU) complex with a number of peptides and identifying polypeptides having a selective binding affinity for the RU complex. The method is based on functional similarities and does not rely on sequence similarities. Prior methods only gave limited success for identifying proteins which contain an SH3 domain due to the minimal sequence homology among known SH3 proteins. It has been found that small peptide RUs in multivalent form have reduced specificity for a given functional domain compared to monomer RUs. Multivalent RU complexes are particularly suited to screening for polypeptides containing functional domains that are similar to, but not identical in sequence to, the original target functional domain. The new method enables proteins having a common function to be identified. Identification of novel SH3 proteins will be useful for a better understanding of cell growth, malignancy, signal transduction processes, etc. New candidate drugs can be identified, and their specificities (e.g. pharmacological activities) can be assessed using the method of the invention.

SO Sequence 2873 BP; 903 A; 506 C; 623 G; 841 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	2873
Score:	1.44e-21	2873
Percent Similarity:	253.00	47
Best Local Similarity:	87.10%	7
Query Match:	75.81%	8
DB:	74.19%	Indels: 0
DB:	2	Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AAT39799 (1-2873)

QY 3 GlnlelaaglnvalilleaSerTyThralathrglyProgluglnleuthleuAla 22

Db 1119 GAGATTGCTCAGGTAACCTCAGCATATGTTGCTTCTGCTGACAACTTACCTTGCA 1178

QY 23 Progluglnleuileuileuileuileuileuileuileuileuileuileuileu 42

Db 1179 CAGGACAGTTAATTAATTTCTAAAGAAAATACAGGTGGTGGCAAGAGTTA 1238

QY 43 GlnlaaArgGlyLysLysArgGlnleuileuileuileuileuileuileuileu 62

Db 1239 CAGGCCAGAGAAAAAGCAGACGAAAGGATGTTCTCTCCAGCATGTTAACTTTTG 1298
 QY 63 SerPro 64
 Db 1299 GGTCCA 1304
 RESULT 14
 AA239011
 ID AA239011 standard; cDNA; 3593 BP.
 AC AA239011;
 XX 28-FEB-2000 (first entry)
 DT Mouse E2e2 coding sequence.
 XX
 DE Mouse; murine; E2e2; endocytosis; vesicular trafficking;
 KW regulation; actin cytoskeleton; detection; cancer; infection;
 KW BH-domain and SH3-domain regulator of endocytosis; anticancer;
 KW antiproliferative; antiviral; s.
 XX
 OS Mus sp.
 PN WO955728-A2.
 XX 04-NOV-1999.
 PD 27-APR-1999; 99WO-CAN00375.
 PF 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 XX Egan SE, Wang W, Sengar A;
 XX WPI; 2000-052802/04.
 DR P-PSDB; AAY57445.
 DR
 XX New nucleic acid encoding E2e2 and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 FT
 XX
 PS Claim 25; Page 46-48; 99pp; English.
 XX
 CC The present sequence encodes mouse E2e2. The present invention
 CC specifically describes mammalian E2e2 and 2 proteins (I) and their splice
 CC variants (Bse = EH-domain and SH3-domain regulator of endocytosis). (I)
 CC are involved in regulation of clathrin-mediated endocytosis (as a complex
 CC with E2e2 protein), vesicular trafficking and actin cytoskeleton.
 CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
 CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
 CC polynucleotide; agents that downregulate expression of E2e2 genes or
 CC antagonists of an E2e2 binding partner are used to treat diseases
 CC associated with undesirable endocytosis and resulting changes in cellular
 CC function. Particularly overexpression of E2e2 is used to block clathrin-
 CC mediated endocytosis in vivo or in cell cultures, while administration of
 CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of
 CC (I) or Ab are used to suppress abnormal proliferation of cells that can
 CC be stimulated to proliferate by a growth factor receptor; and similar
 CC compounds (also inactive E2e2 mutants) can be used to prevent viral
 CC infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an E2e2-E2e2 complex, then binding dynamin to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission
 CC
 XX
 SQ Sequence 3593 BP; 1171 A; 783 C; 850 G; 789 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,9e-21 Length: 3593
 Score: 253.00 Matches: 47
 Percent Similarity: 87.10% Conservative: 7

Best Local Similarity: 75.81% Mismatches: 8
 Query Match: 74.19% Indels: 0
 DB: 3 Gaps:
 US-09-720-934-2_COPY_999_1062 (1-64) x AA239011 (1-3593)
 QY 3 GlnIleAlaGlnValIleAlaSerTYrThrAlaThrGlyProGlnGlnIleuThrIleuAla 22
 Db 3046 GAGATCGCTCAAGTAACCTCAGCATATGCTGCTTCAGGACCTGACAGGCTCAGCCTTGCG 3105
 QY 23 ProGlyGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42
 Db 3106 CCAGGACAGTATATATATATCTTAAAGAAAACAAACCCGGTGGTGGCAAGAGAGCTA 3165
 QY 43 GlnAlaArgGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 62
 Db 3166 CAGGCCAGAGGAGAAAGAAACAGACGAAAGGATGTTCTCTCCAGCATGTTAAAGCTGCTA 3225
 QY 63 SerPro 64
 Db 3226 GGTCCA 3231
 RESULT 15
 ABK43586
 ID ABK43586 standard; cDNA; 3746 BP.
 XX
 AC ABK43586;
 XX 05-JUN-2002 (first entry)
 DT
 XX
 DE DNA encoding novel central nervous system protein #166.
 XX
 KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
 KW adenocarcinoma; reproductive system disorder; testicular feminisation;
 KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;
 KW myocardial infarction; wound healing; cell proliferation; skin aging;
 KW food additive; food preservative; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 PN WO200155318-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001332.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 02-MAR-2000; 2000US-0184664P.
 PR 16-MAR-2000; 2000US-0186350P.
 PR 17-MAR-2000; 2000US-0189874P.
 PR 18-APR-2000; 2000US-0190076P.
 PR 19-MAY-2000; 2000US-0205151P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216447P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.

PR	14-AUG-2000	2000US-02252666
PR	14-AUG-2000	2000US-02252678
PR	14-AUG-2000	2000US-02252687
PR	14-AUG-2000	2000US-02252700
PR	14-AUG-2000	2000US-02254472
PR	14-AUG-2000	2000US-02255757
PR	14-AUG-2000	2000US-02257588
PR	14-AUG-2000	2000US-02257599
PR	14-AUG-2000	2000US-02262619
PR	14-AUG-2000	2000US-02266811
PR	22-AUG-2000	2000US-02266868
PR	22-AUG-2000	2000US-02271822
PR	30-AUG-2000	2000US-02270092
PR	01-SEP-2000	2000US-02289242
PR	01-SEP-2000	2000US-02292878
PR	01-SEP-2000	2000US-02293433
PR	01-SEP-2000	2000US-02293447
PR	01-SEP-2000	2000US-02293452
PR	05-SEP-2000	2000US-02295039
PR	05-SEP-2000	2000US-02295137
PR	06-SEP-2000	2000US-02304378
PR	06-SEP-2000	2000US-02304382
PR	08-SEP-2000	2000US-02312422
PR	08-SEP-2000	2000US-02312432
PR	08-SEP-2000	2000US-02312442
PR	08-SEP-2000	2000US-02314132
PR	08-SEP-2000	2000US-02314142
PR	08-SEP-2000	2000US-02320802
PR	12-SEP-2000	2000US-02320812
PR	12-SEP-2000	2000US-02319682
PR	14-SEP-2000	2000US-02323272
PR	14-SEP-2000	2000US-02323282
PR	14-SEP-2000	2000US-02323292
PR	14-SEP-2000	2000US-02324002
PR	14-SEP-2000	2000US-02324012
PR	14-SEP-2000	2000US-02330632
PR	14-SEP-2000	2000US-02330642
PR	14-SEP-2000	2000US-02330652
PR	21-SEP-2000	2000US-02342232
PR	21-SEP-2000	2000US-02342742
PR	25-SEP-2000	2000US-02349272
PR	25-SEP-2000	2000US-02349282
PR	25-SEP-2000	2000US-02349292
PR	25-SEP-2000	2000US-02349842
PR	27-SEP-2000	2000US-02358342
PR	27-SEP-2000	2000US-02358352
PR	29-SEP-2000	2000US-02363272
PR	29-SEP-2000	2000US-02363672
PR	29-SEP-2000	2000US-02363682
PR	29-SEP-2000	2000US-02363692
PR	29-SEP-2000	2000US-02363702
PR	02-OCT-2000	2000US-02368022
PR	02-OCT-2000	2000US-02370372
PR	02-OCT-2000	2000US-02370382
PR	02-OCT-2000	2000US-02370392
PR	02-OCT-2000	2000US-02370402
PR	13-OCT-2000	2000US-02393932
PR	13-OCT-2000	2000US-02393932
PR	20-OCT-2000	2000US-02409602
PR	20-OCT-2000	2000US-02411222
PR	20-OCT-2000	2000US-02411782
PR	20-OCT-2000	2000US-02411782
PR	20-OCT-2000	2000US-02411782
PR	20-OCT-2000	2000US-02418082
PR	20-OCT-2000	2000US-02418092
PR	01-NOV-2000	2000US-02446172
PR	01-NOV-2000	2000US-02446172
PR	08-NOV-2000	2000US-02464752
PR	08-NOV-2000	2000US-02464762
PR	08-NOV-2000	2000US-02464772
PR	08-NOV-2000	2000US-02464782
PR	08-NOV-2000	2000US-02465342
PR	08-NOV-2000	2000US-02465342

PR	08-NOV-2000	2000US-0246525P	
PR	08-NOV-2000	2000US-0246526P	
PR	08-NOV-2000	2000US-0246527P	
PR	08-NOV-2000	2000US-0246528P	
PR	08-NOV-2000	2000US-0246532P	
PR	08-NOV-2000	2000US-0246509P	
PR	08-NOV-2000	2000US-0246610P	
PR	08-NOV-2000	2000US-0246611P	
PR	17-NOV-2000	2000US-0246613P	
PR	17-NOV-2000	2000US-0249207P	
PR	17-NOV-2000	2000US-0249208P	
PR	17-NOV-2000	2000US-0249209P	
PR	17-NOV-2000	2000US-0249210P	
PR	17-NOV-2000	2000US-0249211P	
PR	17-NOV-2000	2000US-0249212P	
PR	17-NOV-2000	2000US-0249213P	
PR	17-NOV-2000	2000US-0249214P	
PR	17-NOV-2000	2000US-0249215P	
PR	17-NOV-2000	2000US-0249216P	
PR	17-NOV-2000	2000US-0249217P	
PR	17-NOV-2000	2000US-0249218P	
PR	17-NOV-2000	2000US-0249244P	
PR	17-NOV-2000	2000US-0249245P	
PR	17-NOV-2000	2000US-0249246P	
PR	17-NOV-2000	2000US-0249256P	
PR	17-NOV-2000	2000US-0249257P	
PR	17-NOV-2000	2000US-0249259P	
PR	01-DEC-2000	2000US-0250160P	
PR	01-DEC-2000	2000US-0250161P	
PR	05-DEC-2000	2000US-0251030P	
PR	05-DEC-2000	2000US-0251030P	
PR	05-DEC-2000	2000US-0251988P	
PR	05-DEC-2000	2000US-0256719P	
PR	06-DEC-2000	2000US-0251479P	
PR	08-DEC-2000	2000US-0251856P	
PR	08-DEC-2000	2000US-0251868P	
PR	08-DEC-2000	2000US-0251869P	
PR	08-DEC-2000	2000US-0251989P	
PR	08-DEC-2000	2000US-0251990P	
PR	11-DEC-2000	2000US-0254097P	
PR	05-JAN-2001	2001US-0259678P	
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Rosen CA, Barash SC, Ruben SM;		
DR	WPI: 2001-581633/65.		
DR	P-PSDB; AAU07256.		
XX			
PT	New isolated nucleic acid encoding a protein for diagnosing, preventing,		
PT	treating or ameliorating medical conditions and used as food additives or		
XX	preservatives.		
XX			
XX	Claim 1; SEQ ID NO 176; 837p; English.		
XX			
XX	The invention describes an isolated nucleic acid molecule (I) encoding a		
CC	novel central nervous system protein. (I) and polypeptides (Iii) encoded		
CC	by (I), are used to treat a medical conditions and in diagnosis of a		
CC	pathological condition. Disorders which are diagnosed or treated include		
CC	autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative		
CC	disorders e.g. neoplasms of the breast or liver, cardiovascular disorders		
CC	e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,		
CC	angiogenesis, nervous system disorders e.g. Alzheimer's disease and		
CC	amyloidotic lateral sclerosis, infections caused by bacteria, viruses		
CC	e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders		
CC	e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,		
CC	adenocarcinomas and irritable bowel syndrome, reproductive system		
CC	disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes		
CC	and pituitary dwarfism, cancers and disorders at the cellular level e.g.		
CC	leukemia, disorders involving neovascularisation e.g. malignancies,		
CC	respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.		
CC	acute kidney failure and blood related disorders e.g. myocardial		
CC	infarction. The polypeptides can also be used to aid wound healing and		

CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Alignment Scores:

Pred. No.: 2e-21
 Score: 253.00
 Percent Similarity: 87.10%
 Best Local Similarity: 75.81%
 Query Match: 74.19%
 DB: 4

Length: 3746
 Matches: 47
 Conservative: 7
 Mismatches: 8
 Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x ABK43586 (1-3746)

QY	3	GIHIIeAIIaGIIIVaIIIIeAIIaSaRTYrThrAlaThrGIYProGIuGIuLeuThrLeuAla	22
DB	1045	GAGATTGCTCAGGTAACCTTCAGGCAATATGCTTCTGTTCTGAACTTAGCCTTGCA	1104
QY	23	ProGIYGIuLeuIIeLeuIIeAIIaGIIIVaIIIIeAIIaSaRTYrThrAlaThrGIYProGIuGIuLeu	42
DB	1105	CCAGGACAGTTATATTAATTCCTAAGAAAAATACAAGTGGTGTGGCAAGAGAGCTTA	1164
QY	43	GIHIIaIaIrgIYIYsIYsAIIaGIIIVaIIIIeAIIaSaRTYrThrAlaThrGIYProGIuGIuLeu	62
DB	1165	CAGGCCAGAGAAAAAGCGACAGAAAGAGTGTTCCTGCAAGTCAAGTTAACTTTTG	1224
QY	63	SerPro 64	
DB	1225	GGTCCA 1230	

Search completed: July 1, 2004, 13:32:04
 Job time : 205.546 secs

PA (UYN-) UNIV NORTH CAROLINA.
 XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, Mcconnell SJ,
 PI WPI; 1996-465045/46.
 DR P-PSDB; AAM05395.
 XX Identifying polypeptide(s) having specific functional domain (esp. SH3
 PT domain) - comprises detecting selective binding to recognition unit,
 PT regardless of sequence homology
 XX
 PS Claim 76; Fig 50; 174pp; English.
 XX AAT39786-T39803 represent novel human and mouse genes encoding Src-
 CC homology region 3 (SH3) domain containing proteins that can be used in
 CC the method of the invention. SH3 domain containing proteins play a role
 CC in signalling and structural elements of cells. The method of the
 CC invention is for identifying polypeptides containing functional domains
 CC of interest (especially SH3 domains). The method comprises contacting a
 CC multivalent recognition unit (RU) complex with a number of peptides and
 CC identifying polypeptides having a selective binding affinity for the RU
 CC complex. The method is based on functional similarities and does not rely
 CC on sequence similarities. Prior methods only gave limited success for
 CC identifying proteins which contain an SH3 domain due to the minimal
 CC sequence homology among known SH3 proteins. It has been found that small
 CC peptide RUs in multivalent form have reduced specificity for a given
 CC functional domain compared to monomer RUs. Multivalent RU complexes are
 CC particularly suited to screening for polypeptides containing functional
 CC domains that are similar to, but not identical in sequence to, the
 CC original target functional domain. The new method enables proteins having
 CC a common function to be identified. Identification of novel SH3 proteins
 CC will be useful for a better understanding of cell growth, malignancy,
 CC signal transduction processes, etc. New candidate drugs can be
 CC identified, and their specificities (e.g. pharmacological activities) can
 CC be assessed using the method of the invention
 XX
 SQ Sequence 1389 BP; 404 A; 370 C; 357 G; 258 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6.04e-37 Length: 1389
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AAT39795 (1-1389)
 QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrrProTrrPARGAlaLysAspAsnHisIleu 20
 Db 676 GTGAGGGGCTACCAAGCTCAAGCCCTATATCCTTGAGAGCCAAAGAACCAACCACTTA 735
 QY 21 AsnPhaAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTrpTrpPheGly 40
 Db 736 AATTTTAACAATAATATGTATCTATCAACCGTCCTGGAAACACAGACATGTGGTGTGGA 795
 QY 41 GluValGlnGlyGlnLysGlyTrpPheProLysSerTrpValLysLeuIleSerGly 59
 Db 796 GAAGTTCAAGGTCAGAAAGGTTGGTTCCTCCCAAGTCCTTACGTGAACCTCAATTCAGGG 852
 RESULT 2
 AA234574
 ID AA234574 standard; cDNA; 3231 BP.
 XX
 AC AA234574;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 9.
 XX
 XX SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KM

KM haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 2..1927
 FT /*tag= a
 XX
 PN W09953062-A2.
 XX
 PD 21-OCT-1999.
 PD
 XX
 PF 16-APR-1999; 99MO-US008371.
 PF
 XX 16-APR-1998; 98US-0082007P.
 PR
 XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 PA
 XX Korenberg JR, Chen X;
 PI WPI; 1999-633829/54.
 DR P-PSDB; AAY32158.
 DR
 XX Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 PT
 PS Claim 2; Fig 14; 99pp; English.
 XX
 XX This is the nucleotide sequence of a non-full-length cDNA (clone 9)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality; myeloproliferative
 CC disorder; haematopoietic disorder; platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 3231 BP; 1004 A; 721 C; 712 G; 794 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.8e-36 Length: 3231
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AA234574 (1-3231)
 QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrrProTrrPARGAlaLysAspAsnHisIleu 20
 Db 1214 GTGAGGGGCTACCAAGCTCAAGCCCTATATCCTTGAGAGCCAAAGAACCAACCACTTA 1273
 QY 21 AsnPhaAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTrpTrpPheGly 40
 Db 1274 AATTTTAACAATAATATGTATCTATCAACCGTCCTGGAAACACAGACATGTGGTGTGGA 1333

QY 41 GluValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuIleSerGly 59
 DB 1334 GAAGTTCAGGCTCAGAGGCTGTGCTCCCAAGCTTACGTAACACTATTTCAGGG 1390
 RESULT 3
 AA234572
 ID AA234572 standard; cDNA; 5195 BP.
 AC AA234572;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 11.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 239..3886
 FT /*tag= a
 PN WO9953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WO-US008371.
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI; 1999-633829/54.
 DR P-PSDB; AAY32156.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 10; 99p; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukaemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC association of gains in chromosome 21, low platelets in deletion for 21,
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;
 Alignment Scores:

Pred. No.: 3.33e-36 Length: 5195
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-720-934-2_copy_908_966 (1-59) x AA234572 (1-5195)
 QY 1 ValGlnGlyLeuGlnAlaGlnAlaLeuTyrProTyrPheArgAlaLysLysAspAsnHisLeu 20
 DB 2960 GTGAGGGGGCTACAGGCTCAAGCCCTATATCTTGGAGAGCCAAAAAGCAACCACTTA 3019
 QY 21 AsnPheAsnLysAsnAspValIleThrValLeuGlnGlnLysMetTyrTrpPheGly 40
 DB 3020 AATTTTACAAAATGATGTCATCACCGCTCTGGAAACGACAGACATGAGTGTGGA 3079
 QY 41 GluValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuIleSerGly 59
 DB 3080 GAAGTTCAGGCTCAGAGGCTGTGCTCCCAAGCTTACGTAACACTATTTCAGGG 3136
 RESULT 4
 AA234570
 ID AA234570 standard; cDNA; 5195 BP.
 AC AA234570;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 208..3642
 FT /*tag= a
 PN WO9953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WO-US008371.
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI; 1999-633829/54.
 DR P-PSDB; AAY32154.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 5; 99p; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA corresponding to a
 CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the
 CC development of platelets and the pathogenesis of leukaemias, both in
 CC general and in particular those involving the megakaryocytic lineage. The
 CC SH3D1A gene maps to the small candidate region for low platelets on
 CC chromosome 21. Sequencing of 5 different sizes of cDNA clone (see
 CC AA234570-74) suggests that at least 3 isoforms exist. The invention
 CC provides methods for the diagnosis and treatment of megakaryocytic
 CC abnormality, myeloproliferative disorder, platelet disorder, acute
 CC leukaemia, neural disorders, thrombocytopenia, platelet disorder on

CC chromosome 21, low platelets in deletion for 21, association of gains in
CC chromosome 21 with leukemias, neural abnormalities, dysfunction and
CC disorders including brain malformations and corresponding cognitive
CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
CC also provided for: suppressing cells unable to regulate themselves;
CC screening for a somatic alteration in the SH3D1A gene; monitoring the
CC progress and adequacy of a treatment; monitoring tumor risk progress or
CC megakaryocytic abnormality, myeloproliferative disorder, haematopoietic
CC disorder, platelet disorder or leukaemia; and treatment of a subject
CC (including a prenatal subject) having megakaryocytic abnormality,
CC myeloproliferative disorder, platelet disorder, leukaemia or neural
CC disorder using a nucleic acid that expresses SH3D1A or its antisense
CC nucleic acid

XX
XX
SQ Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.34e-36 Length: 5199
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AA234570 (1-5199)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTPrAArgAlaLysLysAspAsnHisLeu 20
Db 2929 GTGAGGGGGCTACAGGCTCAAGCCCTATATCTTGGAGAGCCCAAAAAGACACCACTTA 2998

QY 21 AsnPheAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTPrPheGly 40
Db 2989 AATTTTAACAAAAGATGATCATCAACCGCTCTGGAACGACAGCATGTGGTTTGA 3048

QY 41 GluValGlnGlyGlnLysGlyTPrPheProLysSerTyValLysLeuLisSerGly 59
Db 3049 GAAGTTCAAGGTCAAGAGGTTGGTCCCAAGCTTACGTGAACACTATTTCAGGG 3105

RESULT 5
AA234571
ID AA234571 standard; cDNA; 5458 BP.

XX
XX
AC AA234571;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA clone 21.

XX
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.

XX
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 267..3929
FT /tag= a
XX
XX
PN W09953062-A2.

XX
XX
PD 21-OCT-1999.
XX
XX
PF 16-APR-1999; 99WO-US008371.
XX
XX
PR 16-APR-1999; 98US-0082007P.
XX
XX
PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
XX
PI Korenberg JR, Chen X;
XX
XX
DR WPI, 1999-633829/54.

DR P-PSDB; AA232155.
XX
XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
XX diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
XX
PS Claim 2; Fig 8; 99p; English.

CC This is the nucleotide sequence of full-length cDNA (clone 21)
CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
CC contributes to the development of platelets and the pathogenesis of
CC leukemias, both in general and in particular those involving the
CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
CC region for low platelets on chromosome 21. Sequencing of 5 different
CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
CC least 3 isoforms exist. The invention provides methods for the diagnosis
CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
CC association of gains in chromosome 21 with leukemias, neural
CC abnormalities, dysfunction and disorders including brain malformations
CC and corresponding cognitive dysfunction, microcephaly, lissencephaly,
CC and colpocephaly. Methods are also provided for: suppressing cells unable
CC to regulate themselves; screening for a somatic alteration in the SH3D1A
CC gene; monitoring the progress and adequacy of a treatment; monitoring
CC tumor risk progress or megakaryocytic abnormality, myeloproliferative
CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
CC treatment of a subject (including a prenatal subject) having
CC megakaryocytic abnormality, myeloproliferative disorder, platelet
CC disorder, leukaemia or neural disorder using a nucleic acid that
CC expresses SH3D1A or its antisense nucleic acid

XX
XX
SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.55e-36 Length: 5458
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AA234571 (1-5458)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTPrAArgAlaLysLysAspAsnHisLeu 20
Db 3003 GTGAGGGGGCTACAGGCTCAAGCCCTATATCTTGGAGAGCCCAAAAAGACACCACTTA 3062

QY 21 AsnPheAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTPrPheGly 40
Db 3063 AATTTTAACAAAAGATGATCATCAACCGCTCTGGAACGACAGCATGTGGTTTGA 3122

QY 41 GluValGlnGlyGlnLysGlyTPrPheProLysSerTyValLysLeuLisSerGly 59
Db 3123 GAAGTTCAAGGTCAAGAGGTTGGTCCCAAGCTTACGTGAACACTATTTCAGGG 3179

RESULT 6
AA239009
ID AA239009 standard; cDNA; 3723 BP.

XX
XX
AC AA239009;
XX
DT 28-FEB-2000 (first entry)
XX
XX
DE Mouse Beal coding sequence.
XX
XX
KW Mouse; murine; Esel; Esel2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.

XX
XX
OS Mus sp.
XX
XX
PN W09955728-A2.

```
XX 04-NOV-1999.
PD
XX
XX 27-APR-1999; 99WO-CA000375.
PF
XX 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
PA (HSCR-) HSC RES & DEV LP.
XX
XX Egan SE, Wang W, Sengar A;
XX WPI, 2000-052802/04.
DR P-PSDB; AAY57444.
XX
XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
PT
XX
PS Claim 6; Page 40-42; 99pp; English.
XX
XX The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (1) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp1s protein), vesicular trafficking and actin cytoskeleton.
CC Generally (1) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (1)-specific antibodies (Ab); sequences antisense to the (1)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (1) or Ab are used to promote endocytosis of selected cells. (Ant)agonists of
CC (1) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1s complex, then binding dynamn to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,13e-35 Length: 3723
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.31% Mismatches: 0
Query Match: 98.48% Indels: 0
DB: 3 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AA239009 (1-3723)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPaGAlaLysLysAspAsnHisLeu 20
DB 2716 GTGGAAGGCGCTACAGCGCAGCCCTGTATCCCTGGAGAGCCAAAAAGACAACACCTTA 2775
QY 21 AsnPhaEnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTyrTTPheGly 40
DB 2776 AATTTTAACAAAGAGACGTATCAACCGTTCTGGAACAGCAAGACATGTGGTGGT 2835
QY 41 GluValGlnGlyGlnLysGlyTTPhePheProLysSerTyrValLysLeuIleSerGly 59
DB 2836 GAAGTTCAGAGTCAGAGAGGTTGGTCCCAAGTCTTACGTGAAGAACTCATTTCAAGG 2892
RESULT 7
AA239008
ID AA239008 standard; cDNA; 5082 BP.
XX
XX AA239008;
XX
XX 28-FEB-2000 (first entry)
DT
```

```
XX Mouse Ese1 full length cDNA sequence.
DE
XX
XX Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX
XX Mus sp.
XX
XX W09955728-A2.
XX
XX 04-NOV-1999.
XX
XX 27-APR-1999; 99WO-CA000375.
XX
XX 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
XX (HSCR-) HSC RES & DEV LP.
XX
XX Egan SE, Wang W, Sengar A;
XX WPI, 2000-052802/04.
DR P-PSDB; AAY57444.
XX
XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
PT
XX
PS Claim 6; Page 38-40; 99pp; English.
XX
XX The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (1) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp1s protein), vesicular trafficking and actin cytoskeleton.
CC Generally (1) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (1)-specific antibodies (Ab); sequences antisense to the (1)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (1) or Ab are used to promote endocytosis of selected cells. (Ant)agonists of
CC (1) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1s complex, then binding dynamn to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,69e-35 Length: 5082
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.31% Mismatches: 0
Query Match: 98.48% Indels: 0
DB: 3 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AA239008 (1-5082)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPaGAlaLysLysAspAsnHisLeu 20
DB 2973 GTGGAAGGCGCTACAGCGCAGCCCTGTATCCCTGGAGAGCCAAAAAGACAACACCTTA 3032
QY 21 AsnPhaEnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTyrTTPheGly 40
DB 3033 AATTTTAACAAAGAGACGTATCAACCGTTCTGGAACAGCAAGACATGTGGTGGT 3092
```

QY 41 GluValGlnGlyGlnIysGlyTTPheProIysSerTyrValIlystLeuIleSerGly 59
 DB 3093 GAAGTTCAGAGTCAGAAAGGGTTGTTCCCAAGCTTACGTGAACCTATTCAAGG 3149

RESULT 8
 ID AAZ39025 standard; cDNA; 5144 BP.
 AC AAZ39025;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Mouse Ese1L coding sequence.
 XX
 KM Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
 KM regulation; actin cytoskeleton; detection; cancer; infection;
 KM EH-domain and SH3-domain regulator of endocytosis; anticancer;
 KM antiproliferative; antiviral; ss.
 XX
 OS Mus sp.
 OS WO955728-A2.
 XX
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO-CA000375.
 XX
 PR 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 XX
 PI Egan SE, Wang W, Sengar A;
 XX WPI: 2000-052802/04.
 DR P-PSDB; AAY57449.
 DR
 PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 PS Claim 6; Page 59-62; 99pp; English.

XX The present invention specifically describes mammalian Ese1 and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Esp1s protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC expression of Ese genes or antagonists of an Ese binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Ese1 is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) is used to promote endocytosis of selected cells.
 CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Ese mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Ese-Esp1s complex, then binding dynamn to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The
 CC present sequence represents mouse Ese1L coding sequence
 CC
 CC Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;
 CC
 CC Alignment Scores:
 CC Pred. No.: 1,71e-35
 CC Score: 323.00
 CC Percent Similarity: 100.00%
 CC Best Local Similarity: 98.31%
 CC Query Match: 98.48%
 CC
 CC DB: 3
 CC Length: 5144
 CC Matches: 58
 CC Conservative: 1
 CC Mismatches: 0
 CC Indels: 0
 CC Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AAZ39025 (1-5144)

QY 1 ValGlnGlyLeuGlnIleGlnIleValIleuGlnIleuGlnIleuGlnIleuGlnIleu 20
 DB 2716 GTGGAAGGGCTTAACAGCGCAAGCCCTGTATCCCTGGAGAACCAAAAAGCAACACTTA 2775

QY 21 AsnPhaAsnIysAsnAspValIleIleValIleuGlnIleuGlnIleuGlnIleuGlnIleu 40
 DB 2776 AATTTTACAAAGTAGCGTCATCCGTTCTGGAACAGCAAGCATGTGTGTTGGA 2835

QY 41 GluValGlnGlyGlnIysGlyTTPheProIysSerTyrValIlystLeuIleSerGly 59
 DB 2836 GAAGTTCAGAGTCAGAAAGGGTTGTTCCCAAGCTTACGTGAACCTATTCAAGG 2892

RESULT 9
 ID AAZ39024 standard; cDNA; 5738 BP.
 AC AAZ39024;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Mouse Ese1L cDNA sequence.
 XX
 KM Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
 KM regulation; actin cytoskeleton; detection; cancer; infection;
 KM EH-domain and SH3-domain regulator of endocytosis; anticancer;
 KM antiproliferative; antiviral; ss.
 XX
 OS Mus sp.
 OS WO955728-A2.
 XX
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO-CA000375.
 XX
 PR 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 XX
 PI Egan SE, Wang W, Sengar A;
 XX WPI: 2000-052802/04.
 DR P-PSDB; AAY57449.
 DR
 PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 PS Claim 6; Page 56-59; 99pp; English.

XX The present invention specifically describes mammalian Ese1 and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Esp1s protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC expression of Ese genes or antagonists of an Ese binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Ese1 is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) is used to promote endocytosis of selected cells.
 CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Ese mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Ese-Esp1s complex, then binding dynamn to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The
 CC present sequence represents mouse Ese1L coding sequence
 CC

CC present sequence represents mouse Escl1 cDNA sequence
XX Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
SQ Alignment Scores:
Pred. No.: 1.97e-35 Length: 5738
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.31% Mismatches: 0
Query Match: 98.48% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AA239024 (1-5738)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTPrArgAlaLysLysAspAsnHisLeu 20
Db 2975 GTGGAGGGGCTCAAGGCGCAAGCCCTGTATCCCTGGAGGCCCAAAAAAGCAACCACTTA 3034
QY 21 AsnPhenAlaAspValIleThrValLeuGlnGlnAspMetTPTTPPhgI 40
Db 3035 AATTATAACAAAGTGAAGTCATCACCGTTCGGAAACGACACATGTGTGTTTGA 3094
QY 41 GluValGlnGlyGlnLysGlyTYrPheProLysSerTYrValLysLeuIleSerGly 59
Db 3095 GAAGTTCAAGGTCAAGAGGGTGTGTTCCCAAGTCTTAAGTCAATTCATTCAGGG 3151
RESULT 10
AAS84763
ID AAS84763 standard; cDNA; 7435 BP.
XX AAS84763;
XX AC
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #20567.
XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PS 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI, 2001-639362/73.
XX DR P-PsDB; ABG20576.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 20567; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (II) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting (II) as a
XX CC polypeptide in tissue, as molecular weight markers and as a food

XX supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptides and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;
XX Alignment Scores:
Pred. No.: 3.87e-34 Length: 7435
Score: 315.00 Matches: 59
Percent Similarity: 98.33% Conservative: 0
Best Local Similarity: 98.33% Mismatches: 0
Query Match: 96.04% Indels: 1
DB: Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AAS84763 (1-7435)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTPrArgAlaLysLysAspAsnHisLe 20
Db 3019 GTGGAGGGGCTCAAGGCGCAAGCCCTATATCTTGGAGGCCCAAAAAAGCAACCACTT 3078
QY 20 uAsnPhenAlaAspValIleThrValLeuGlnGlnAspMetTPTTPPhgI 40
Db 3079 AATTATAACAAAGTGAAGTCATCACCGTTCGGAAACGACACATGTGTGTTTGG 3138
QY 40 yGluValGlnGlyGlnLysGlyTYrPheProLysSerTYrValLysLeuIleSerGly 59
Db 3139 AGAAGTTCAAGGTCAAGAGGGTGTGTTCCCAAGTCTTAAGTCAATTCATTCAGGG 3196
RESULT 11
AAS84759
ID AAS84759 standard; cDNA; 395 BP.
XX AAS84759;
XX AC
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #20563.
XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PS 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI, 2001-639362/73.
XX DR P-PsDB; ABG20572.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX CC

PS Claim 1; SEQ ID NO 20563; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic

CC coding sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 395 BP; 114 A; 96 C; 100 G; 85 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,91e-31 Length: 395

Score: 280.00 Matches: 49

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 85.37% Indels: 0

DB: 5 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AAS84759 (1-395)

QY 11 ProTPrATgAlaLySLySAspAAsnHisLeuVal 30

Db 1 CCTTGGAGAGGCCAAAAGACACACCTTAATTATACAAATGATGTCACCGTC 60

QY 31 LeuGluGInGInAspMetTTrpPheGlyGluValGInGInLySgLyTTrpPhePro 50

Db 61 CTGGAACGACAGACATGTGTGAGTTCGAGAGAGTTCAAGGTCAAGAGGTTGTTCCCC 120

QY 51 LySserTyValLyLeuIleSergLy 59

Db 121 AAGTCTTACGTGAACATCATTTCAAGGG 147

RESULT 12

ABA66139

ID ABA66139 standard; DNA; 206 BP.

XX

AC ABA66139;

XX

DT 01-FEB-2002 (first entry)

XX

DE Human foetal liver single exon nucleic acid probe #14444.

XX

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX

OS Homo sapiens.

XX

PN WO200157277-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000669.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR,

XX

DR WPI; 2001-483447/52.

XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human foetal liver.

XX

XX

PS Claim 4; SEQ ID NO 14444; 639pp + Sequence Listing; English.

XX

XX The invention relates to a single exon nucleic acid probe for measuring

CC human gene expression in a sample derived from human foetal liver. The

CC single exon nucleic acid probes may be used for predicting, measuring and

CC displaying gene expression in samples derived from human foetal liver. The

CC present sequence is a single exon nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 206 BP; 69 A; 33 C; 55 G; 49 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,48e-27 Length: 206

Score: 251.00 Matches: 44

Percent Similarity: 86.44% Conservative: 7

Best Local Similarity: 74.58% Mismatches: 8

Query Match: 76.52% Indels: 0

DB: 4 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x ABA66139 (1-206)

QY 1 ValGluGlyLeuGInAlaGInAlaLeuTyProTPrATgAlaLySLySAspAAsnHisLeu 20

Db 10 GTAGAAAACCTTAAAGACACAGGCCCTTGTTCCTGAGTGCAGAAAGATTAACCATTTG 69

QY 21 AsnPheAsnLySAsnAspValIleThrValLeuGInGInAspMetTTrpPheGly 40

Db 70 AACCTCTCAAAACATGACATATTAAGTCTTGGAGCAGCAAGAAATGTGGTGGTGGG 129

QY 41 GluValGInGInLySgLyTTrpPheProLySserTyValLyLeuIleSergLy 59

Db 130 GAGGTCATGAGAGAGAGAGATGTTCCCAATCTTATGTCAAGATCATTCCTGGG 186

RESULT 13

AA146323

ID AA146323 standard; DNA; 206 BP.

XX

AC AA146323;

XX

DT 17-OCT-2001 (first entry)

XX

DE Probe #15009 used to measure gene expression in human placenta sample.

XX

KW Probe; microarray; human; placenta; antenatal diagnosis;

XX

OS genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200157272-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000663.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
XX Claim 25; SEQ ID NO 15009; 654bp; English.
PS
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 206 BP; 69 A; 33 C; 55 G; 49 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5,48e-27 Length: 206
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
DB: 4 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AA16323 (1-206)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTPrpArgAlaLysLysAspAsnHisLeu 20
Db 10 GTAGAAACTTAAAGCAGAGCCCTTTCTTCTGACTGCAAGAAAGATTAACCACTTG 69
QY 21 AsnPheAsnLysAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly 40
Db 70 AACTCTCAAAACATGACATTAATTACTGCTTGAGAGCAGCAAGAAATTTGGTGGG 129
QY 41 GluValGlnGlyGlnLysGlyTrpPheProLysSerTyrrValLysLeuIleSerGly 59
Db 130 GAGGTGCAATGAGAGAGAGATGTTTCCCAATCTTATGTCAGATCATCTCTGGG 186
RESULT 14
ABS39875
ID ABS39875 standard; DNA; 206 BP.
XX
XX ABS39875;
AC
XX
XX 25-FEB-2003 (first entry)
DT
XX
XX Human liver single exon probe, SEQ ID No 14865.
DE
XX
XX Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
XX Homo sapiens.
OS
XX
XX W0200157273-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000664.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
PR

XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
XX Claim 4; SEQ ID NO 14865; 658bp; English.
PS
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC specification for this patent does not appear in the printed
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 206 BP; 69 A; 33 C; 55 G; 49 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5,48e-27 Length: 206
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
DB: 4 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x ABS39875 (1-206)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTPrpArgAlaLysLysAspAsnHisLeu 20
Db 10 GTAGAAACTTAAAGCAGAGCCCTTTCTTCTGACTGCAAGAAAGATTAACCACTTG 69
QY 21 AsnPheAsnLysAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly 40
Db 70 AACTCTCAAAACATGACATTAATTACTGCTTGAGAGCAGCAAGAAATTTGGTGGG 129
QY 41 GluValGlnGlyGlnLysGlyTrpPheProLysSerTyrrValLysLeuIleSerGly 59
Db 130 GAGGTGCAATGAGAGAGAGATGTTTCCCAATCTTATGTCAGATCATCTCTGGG 186
RESULT 15
ABS14329
ID ABS14329 standard; DNA; 206 BP.
XX
XX ABS14329;
AC
XX
XX 19-AUG-2002 (first entry)
DT
XX
XX Human genome-derived single exon ORF from lung SEQ ID No 14320.
DE
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosterosi;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
OS
XX

PN WO200186003-A2.
 XX 15-NOV-2001.
 PD 30-JAN-2001; 2001WO-US000665.
 PF 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 PS Claim 4; SEQ ID NO 14320; 634bp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridize at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned in the specification, or encoded by the
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease, (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe open reading frame of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 206 BP; 69 A; 33 C; 55 G; 49 T; 0 U; 0 Other;

Alignment Scores:

Pre. No.: 5.48e-27
 Score: 251.00
 Percent Similarity: 86.44%
 Best Local Similarity: 74.58%
 Query Match: 76.52%
 DB: 6

Length: 206
 Matches: 44
 Conservative: 7
 Mismatches: 8
 Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x ABS14329 (1-206)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaLysAspAsnHisLeu 20
 Db 10 GTAGAAACTTAAAGACAGGCCCTTGTCTTGAGCTGCAAGAAAGATTAACCTTG 69
 QY 21 AsnPeAsnLysAsnAspValIleThrValLeuGlnGlnAspMetTTPpGly 40
 Db 70 AACTTCTCAAAACATGACATTTACTGTCTTGAGACAGCAAAATGGTGTGG 129
 QY 41 GluValGlnGlyGlnLysGlyTTPpPheProLysSerTyrValLysLeuIleSerGly 59
 Db 130 GAGGTGCATGAGAGAAAGAGATGTTCCTCAATCTTATGTCAAGATCATTCCTGG 186

Search completed: July 1, 2004, 13:31:56
 Job time : 191.112 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 1391.43 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_740_800
Perfect score: 333

Sequence: 1 VKVYVRYALYPFESRSHDEI.....GGELKGTGWFPANVAKIP 61

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=cmr2.1/USPTO.spool.p/US09720934/runat_30062004_064539_13430/app_query.fasta_1.1386
-DB=GenEmbl -QFMT=fastap -SUFFIX=5ge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNIT3-bits -START=1 -END=1 -MATRIX=blosome2 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEASize=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934 @CNC 1.1 7509 @runat_30062004_064539_13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONELOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*

1:	gb_ba:*
2:	gb_htg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_ph:*
7:	gb_pat:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vt:*
15:	em_da:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rnd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htg_hum:*

40: em_htg_mus:*

41: em_htg_others:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	333	100.0	306	9	AY254341	AY254341 Homo sapi
2	333	100.0	3231	6	BD205037	BD205037 Isolated
3	333	100.0	3241	9	HSU61166	U61166 Human SH3 d
4	333	100.0	3723	10	AF132478	AF132478 Mus muscu
5	333	100.0	5145	10	AF132481	AF132481 Mus muscu
6	333	100.0	5195	6	BD205035	BD205035 Isolated
7	333	100.0	5199	6	BD205033	BD205033 Isolated
8	324	97.3	4321	6	BSM86384	BSM86384 Homo sapi
9	320.5	96.2	475	10	AY127576	AY127576 Mus muscu
10	320.5	96.2	5287	9	AF064243	AF064243 Homo sapi
11	320.5	96.2	5381	9	AF114488	AF114488 Homo sapi
12	320.5	96.2	5458	6	BD205034	BD205034 Isolated
13	320.5	96.2	6439	9	AF114487	AF114487 Homo sapi
14	320.5	96.2	7247	9	AF064244	AF064244 Homo sapi
15	317	95.2	4103	5	AF032118	AF032118 Xenopus l
16	309.5	92.9	3812	10	AF132672	AF132672 Rattus no
17	309.5	92.9	4025	10	AF127798	AF127798 Rattus no
18	216	64.9	3594	10	AF132480	AF132480 Mus muscu
19	216	64.9	4447	6	AX428859	AX428859 Sequence
20	216	64.9	4557	9	AF182199	AF182199 Homo sapi
21	216	64.9	4977	10	AF132479	AF132479 Mus muscu
22	216	64.9	5804	10	AK122480	AK122480 Mus muscu
23	216	64.9	5828	6	BD167848	BD167848 Method fo
24	216	64.9	5828	9	AF248540	AF248540 Homo sapi
25	216	64.9	5938	9	AB033082	AB033082 Homo sapi
26	216	64.9	6091	9	AF182198	AF182198 Homo sapi
27	215	64.6	2873	6	AR175271	AR175271 Sequence
28	186	55.9	169774	2	AC138787	AC138787 Sus scrofa
29	186	55.9	175319	10	AC134483	AC134483 Mus muscu
30	186	55.9	181782	2	AC116970	AC116970 Rattus no
31	186	55.9	217131	10	AC126053	AC126053 Rattus no
32	186	55.9	278501	2	AC098552	AC098552 Rattus no
33	186	55.9	279148	3	AC123507	AC123507 Rattus no
34	184	55.3	2408	3	AF069517	AF069517 Drosophila
35	184	55.3	3750	3	AF054612	AF054612 Drosophila
36	184	55.3	3999	3	AF053957	AF053957 Drosophila
37	184	55.3	59264	2	AC018278	AC018278 Drosophila
38	184	55.3	127035	3	AC006574	AC006574 Drosophila
39	184	55.3	160282	3	AC009253	AC009253 Drosophila
40	184	55.3	273995	3	AE003669	AE003669 Drosophila
41	183	55.0	142974	2	AC129075	AC129075 Felle cat
42	183	55.0	178801	2	AC121075	AC121075 Canis fam
43	183	55.0	184526	2	AC144367	AC144367 Papio ann
44	183	55.0	194157	2	AC144368	AC144368 Gallus ga
45	183	55.0	203423	2	AC137544	AC137544 Gallus ga

RESULT 1

ALIGNMENTS

AY254341
LOCUS Homo sapiens intersec11 isoform 5 (ITSN1) mRNA, partial cds,
DEFINITION alternatively spliced.
ACCESSION AY254341
VERSION AY254341.1 GI:30060229
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Teyba, L., Skrypkina, I., Nikolaenko, A., Fortna, A., Rynditch, A. and
Gardiner, K.
TITLE Tissue-specific regulation of intersec11 transcriptional
isoforms in human and mouse
JOURNAL
REFERENCE 2 (bases 1 to 306)
AUTHORS Teyba, L., Skrypkina, I., Nikolaenko, A., Fortna, A., Rynditch, A. and
Gardiner, K.
TITLE Direct Submision
JOURNAL Submitted (12-MAR-2003) Department of Molecular Oncogenetics,
Institute of Molecular Biology and Genetics of National Academy of
Sciences of Ukraine, 150 Zabolotnogo Str., Kyiv 03143, Ukraine
FEATURES
source
1. .306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="kidney"
/dev_stage="12 week-old embryo"
<1. .306
/gene="ITSN1"
<1. .306
/gene="ITSN1"
/note="alternatively spliced"
/codon_start=1
/product="intersec11 isoform 5"
/protein_id="AAP13099.1"
/db_xref="GI:30060230"
/translation="QHOEPAPKPAVOAPWSTAEKGPLTISAQENVKVVYRYALYPPESR
SHDRTITQPDIVMDESQTGPENLGGELIKGKTGMFPANAYAEKIPENEVAPVKEPTV
H"
ORIGIN
Alignment Scores: 3,016-34 Length: 306
Pred. No.: 333.00 Matches: 61
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 9 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x AY254341 (1-306)
QY 1 VALLYSVALVATYTYRARGAALAEUYYRPROPEGLUSERARSGERHISAPGLIULIE 20
DB 88 GTAAAGAGTGCTATTATACCGGCACTGTACCCCTTGATCCAGAACCATGATGTAATC 147
QY 21 THRILEGIPROGLYASPILEVALMETVALAEPGLUSERGINTHGLYGLIUPROGLYTP 40
DB 148 ACTATCCAGCCAGACAGACATAGTCATGATGATGAAGCCAGCTGAGAACCCGCTGG 207
QY 41 LEUGLYGLYLLEULYSGLYSTHRGTYTPPHEPROALASNTYRZLAGIULYSILE 60
DB 208 CTTGGAGGAGAAATTAAGGAAGACAGAGGCTGTTCCCTCGCAAACTATGCAAGGAATC 267
QY 61 PRO 61
DB 268 CCA 270

RESULT 2
BD205037
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205037
VERSION BD205037.1 GI:33014807
KEYWORDS JP 2002511267-A/5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 5 16-APR-2002;
COMMENT CEDARS SINAI HEALTH SYSTEM ET AL
OS Homo sapiens (human)
PN JP 2002511267-A/5
PD 16-APR-2002
PF 16-APR-1998 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R KORENBERG, XIAO NING CHEN
PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
PC C1201/68,
PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1. .3231
/organism="Homo sapiens (human)"
Location/Qualifiers
1. .3231
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 3,756-33 Length: 3231
Pred. No.: 333.00 Matches: 61
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 6 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x BD205037 (1-3231)
QY 1 VALLYSVALVATYTYRARGAALAEUYYRPROPEGLUSERARSGERHISAPGLIULIE 20
DB 710 GTAAAGAGTGCTATTATACCGGCACTGTACCCCTTGATCCAGAACCATGATGTAATC 769
QY 21 THRILEGIPROGLYASPILEVALMETVALAEPGLUSERGINTHGLYGLIUPROGLYTP 40
DB 770 ACTATCCAGCCAGACAGACATAGTCATGATGATGAAGCCAAACTGAGAACCCGCTGG 829
QY 41 LEUGLYGLYLLEULYSGLYSTHRGTYTPPHEPROALASNTYRZLAGIULYSILE 60
DB 830 CTTGGAGGAGAAATTAAGGAAGACAGAGGCTGTTCCCTCGCAAACTATGCAAGGAATC 889
QY 61 PRO 61
DB 890 CCA 892
RESULT 3
HSU61166
LOCUS 3241 bp mRNA linear PRI 23-JUL-1996
DEFINITION Human SH3 domain-containing protein SH3P17 mRNA, complete cds.
ACCESSION U61166
VERSION U61166.1 GI:1438932

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Sparke,A.B., Hofman,N.G., McConnell,S.J., Fowlkes,D.M. and Kay,B.K.
TITLE Cloning of ligand targets: systematic isolation of SH3 domain-containing proteins
JOURNAL Nat. Biotechnol. 14 (6), 741-744 (1996)
MEDLINE 98294438
PUBMED 9630982
REFERENCE 2 (bases 1 to 3241)
AUTHORS Pirozzi,G., McConnell,S.J., Uveges,A. and Fowlkes,D.M.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1996) CYTOGEN Corp., 307 College Road East, Princeton, NJ 08540, USA
FEATURES
source Location/Qualifiers
1. 3241
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
37. 1599
/codon_start=1
/product="SH3 domain-containing protein SHP17"
/protein_id="AAC50592.1"
/db_xref="GI:1438933"
/translation="MEAEERLKQKEQERKIIIELEKQEAQRAQERDKQWLPHVOQED
EHORPKLHEEELKREESVKKKDGEEKQKQADLGRLFHQHEPAPVQAPST
AEKGLPTISAQENKVVYRALYPPESRSHDEITIQDGIWVDESQTEPGMLGGEL
KQGTWFPNTAAENKIPENYPAKVPVSTSTAPAKLAREPAPLAYTSSSPSTP
NNWADSSSTPSTSTNEKPEKTDMDMAAQAOPVSTAGQLRQASATPATATGSSSP
VIGQKRVGLQALQALYPAKRNKNDVITVLEQDMMWFGEVQGGGWPEK
SVYKISGPIRSTSDSSSPSLKVASPAKPVVSGEIAQVAVASYATGPEQ
LTLAPQQLILIRKKNPGWMBELQARKKQIGMPANVYKLSPGTSKITPTTEPK
STALAAVCQVIGMDYTAQNDELAFNKGQIINVINKEDPDMWKGVEVNGVGLFPSNY
VLTMDPSQ"

ORIGIN
Alignment Scores:
Pred. No.: 3.76e-33 Length: 3241
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x HSU61166 (1-3241)

QY 1 VallysValValTYTYrArGAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
Db 382 GTAAAGTGTGTATTATCCGGGCACTGTACCCCTTTGATCCAGAGCCATATGAAATC 441
QY 21 ThrllleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTTP 40
Db 442 ACTATCCAGCCAGAGACTACTGATGATGATGAAAGCAAACTGAGCAACCCGGCTGG 501
QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTTPheProAlaAsnTYrAlaGluLysIle 60
Db 502 CTGGAGGAGATTAAGAAAGAAAGACAGGATGTTCCCTGCAAACTATGACAGAAATC 561
QY 61 Pro 61
Db 562 CCA 564

RESULT 4
AF132478 3723 bp mRNA linear ROD 09-MAR-1999
LOCUS AF132478 Mus musculus Eset protein mRNA, complete cds.
DEFINITION AF132478
ACCESSION AF132478.1 GI:4378884
VERSION

KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionahli; Muridae; Murinae; Mus.
AUTHORS Sengat,A.S., Wang,W., Bishay,J., Cohen,S. and Egan,S.E.
TITLE The EH and SH3 domain Eset proteins regulate endocytosis by linking to dynamn and Eps15
JOURNAL EMBO J. 18 (5), 1159-1171 (1999)
MEDLINE 99164083
PUBMED 10064583
REFERENCE 2 (bases 1 to 3723)
AUTHORS Sengat,A.S., Wang,W., Cohen,S., Bishay,J. and Egan,S.E.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-1999) Programs in Cancer & Blood Research/Developmental Biology, The Hospital for Sick Children, 555 University Avenue, Toronto, ON M5G-1X8, Canada
FEATURES
source Location/Qualifiers
1. 3723
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1. 3542
/codon_start=1
/product="Eset protein"
/protein_id="A019746.1"
/db_xref="GI:4378885"
/translation="MAQFPPTPGSGSDWVAITYEERAKHDQQLSLKPIAGRTTGDQA
KNPFQSGHPQPLAQTALADNNNDGRDQVFFSLAMKILIKLIGYQLPSTLPVM
GAPPAVASSAPAFGIGLSMPPLVAAPVPMGSIIVGQSPPLVASVPAAPVAPLAN
GAQPVQPLPAFAHPAPATWPKSSFSRSGSLNKLQKQSPVAVSAPAAEMVAP
QSRRLKROLFNHSDKTMGHLTGQARTILMOSLPOQLAIVMLSDIDQGXLT
BEFILLAMHLIDVAMSGQPLPVLPEPTIPSPFRVRVSGSGMSYSSSSQDRLPEPS
SEDOOPEKKLVPTFEDKRENNPERSGVSLEKROALLQOQREGRLLQLEAEER
KREBOEOAKQOLELEKQLEKQRELEERREKREIRERAKLELRQLEMER
NRQELNQRKEQSTGVLLKARKRTLELELALNDKRLQKLEODINCRILATQRE
LESTNKSRELRLAEITTHLQOQLOESQOMIGRLIPKQILSDQKQVQNSLHRSQSL
LKRALBAKELARQOLREDOEVERETRSKLOEIDVNNQLEKLEIHSQOLQOKQSL
EVARLKKEERAKPEMDOKSRLLHPHDEPAKLTQAWSTTEKQPLTISAESVY
VYVRALYPPESRSHDEITIQDGIWVDESQTEPGMLGGELKQGTWFPNTAAENKIP
ENYVTPPAKVPVSTSTAPAKLAREPAPLAYTSSSPSTPNNWADSSSTPSTSTNEK
KRETNWMTWMAQPSLITVPSAGLQKRSAPFTATATGATGAGGVEGLQALQ
LYPAKKNHNLNFNSKSDVITVLEQDMMWFGEVQGGGWPEKSVYKLSPGTSKITPTTEPK
DTGPESPASLKRVASPAKPAIPGEPIFLAMTYTSSSEGDLTPQCGDITVYTKGCD
MWITGVGPKSGVFPNSYVRLKDSGSGTAGKSGISGKREINQVASTATGPEQTL
APGQILIRKKNPGWMBELQARKKQIGMPANVYKLSPGTSKITPTTEPKTAV
QPAVCQVIGMDYTAQNDELAFNKGQIINVINKEDPDMWKGVEVNGVGLFPSNYKL
TTMDPSQ"

ORIGIN
Alignment Scores:
Pred. No.: 4.36e-33 Length: 3723
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AF132478 (1-3723)

QY 1 VallysValValTYTYrArGAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
Db 2212 GTAAAGTGTGTATTATCCGGGCACTGTACCCCTTTGATCCAGAGCCATATGAAATC 2271
QY 21 ThrllleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTTP 40
Db 2272 ACCATCCAGCCAGAGATYTAGTCACTGATGATGAAAGCCAGACTGAGAGCCAGAGATGG 2331
QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTTPheProAlaAsnTYrAlaGluLysIle 60
Db 2332 CTGGAGGAGAGCTGAAAGGAAAGACGAGATGTTCCCTGCAAACTATGACGAAAGATT 2391

polYA_site 4289
/gene="DKFZp686d17173"

ORIGIN

Alignment Scores:
Pred. No.: 7.55e-32 Length: 4321
Score: 324.00 Matches: 60
Percent Similarity: 98.36% Conservative: 0
Best Local Similarity: 98.36% Mismatches: 1
Query Match: 97.30% Indels: 0
Gaps: 0
DB: 9

US-09-720-934-2_COPY_740_800 (1-61) x HSM806384 (1-4321)

QY 1 VallyeValValYrYrYrAgaAlaLeuTYrProPheGluSerArgSerHisaaPgluIle 20
Db 1785 GTAAAGTGTATATTACCGGACCTGTACCCCTTGAATCCAGAAAGCCATGATGAATC 1844

QY 21 ThrTllegInProGlyaspIleValMetValaspGluSerGlnThrGlyGluProGlyTYr 40
Db 1845 ACTATCCAGCCAGAGACATAGTCATGTGATGAAAGCTAAAGCTGAGAGAACCCGCTGG 1904

QY 41 LeuGlyGluGluLeuysGlyLeuThrGlyTYrPheProAlaAsnTYrAlaGluIle 60
Db 1905 CTGGAGAGAAATTAAAGAAAGACAGGGTGTCTCCCTGCAACTATGACAGAAAAATC 1964

QY 61 Pro 61
Db 1965 CCA 1967

RESULT 9
AY127576 475 bp mRNA linear ROD 01-DEC-2002
LOCUS Mus musculus intersectin isoform 5 (Itsn) mRNA, partial cds;
DEFINITION alternatively spliced.
ACCESSION AY127576
VERSION AY127576.1 GI:25989570
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 475)
JOURNAL Slavov, D., Rynditch, A. and Gardiner, K.
AUTHORS Skripkina, I., Tsyba, L.,
TITLE Novel splicing forms of human and mouse intersectin gene
transcripts
JOURNAL Ukr. Biokhim. Zh. (2002) In press
REFERENCE 2 (bases 1 to 475) In press
AUTHORS Skripkina, I., Tsyba, L. and Rynditch, A.
TITLE Direct Submission
JOURNAL Department of Molecular Oncogenetics,
Institute of Molecular Biology and Genetics of National Academy of
Sciences of Ukraine, 150, Zabolotnogo Str., Kiev 03143, Ukraine

FEATURES
source
1..475
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/cx57BLP1"
/db_xref="taxon:10090"
/chromosome="14"
/tissue_type="brain"
/tissue="brain"
/gene="Itsn"
/gene="Itsn"
/note="alternatively spliced"
/codon_start=1
/product="intersectin isoform 5"
/protein_id="AA02285.1"
/db_xref="GI:25989571"
/translation="MHQEPKALATQAPWSTTEKPLTISAGSEVAVVYRLLYPPES
RSHDRITDGPVIMYKGVWBSQSGEGMIGELKGTGMPFANYPANAKIPENEVPT
PAKPVTDLSAPAPKALARETRAPLPVTSSEPTTPNNMADSSSTWSSSNEKPEP"

ORIGIN

Alignment Scores:
Pred. No.: 2.02e-32 Length: 475
Score: 320.50 Matches: 61
Percent Similarity: 92.42% Conservative: 0
Best Local Similarity: 92.42% Mismatches: 0
Query Match: 96.25% Indels: 5
DB: 10 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x AY127576 (1-475)

QY 1 VallyeValValYrYrYrAgaAlaLeuTYrProPheGluSerArgSerHisaaPgluIle 20
Db 91 GTAAAGTGTATATTACCGAGCGCTGTACCCCTTGAATCCAGAAAGCTCAGATGATC 150

QY 21 ThrTllegInProGlyaspIleValMetValaspGluSerGlnThrGlyGluProGlyTYr 35
Db 151 ACCATCCAGCCAGAGACATAGTCATGTGTTAAAGGGAATGGGTGATGAAAGCCAGACT 210

QY 36 GlyGluProGlyTYrLeuGlyGlyLeuLeuysGlyLeuThrGlyTYrPheProAlaAsn 55
Db 211 GGAGAGCCAGAGATGCTTGGAGAGAGAGCTGAAAGGAAAGCGGATGTCTCCCTGCAAA 270

QY 56 TyrAlaGluIlePro 61
Db 271 TATGCAAAAAAGATTCCA 288

RESULT 10
AF064243 5287 bp mRNA linear PRI 21-NOV-1998
LOCUS Homo sapiens intersectin short form mRNA, complete cds.
DEFINITION AF064243
ACCESSION AF064243.1 GI:3859852
VERSION AF064243.1 GI:3859852
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 5287)
JOURNAL Guipponi, M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and
AUTHORS Antonarakis, S.E.
TITLE Two isoforms of a human intersectin (ITSN) protein are produced by
brain-specific alternative splicing in a stop codon
JOURNAL Genomics 53 (3), 369-376 (1998)
AUTHORS Guipponi, M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) Genetics and Microbiology, CMC, 1 rue
AUTHORS Michel-Serret, Geneva 4 CH-1211, Switzerland
TITLE Location/Qualifiers

FEATURES
source
1..5287
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="fetus"
/codon_start=1
/product="intersectin short form"
/protein_id="AAC78610.1"
/db_xref="GI:3859853"
/translation="MAQPTPPGSGIDWALTVERAKDQGFHSXKPSGPTTGPDA
RNFPQSGLPQPVLAQIWAIDAMNDNGMDQVFSIAMLTKIKQIGYQUPALPBYM
KQOPVASAPFGMGIGASMPPLTAVAPVMSGIPVGMSPPLVSSVPTAAVAPPLAN
GAPVTPQIPAPFAHPAAATLPKSSSFRSGSGSLNKLQKASQFVDAVVPVEMAVP
QSRILKTRLPUNSHDKTMSGHILTPQARTIMQSSILPQALASILNLSIDIDQGLTA
EESTLAMIHIDVAMSGQPLPVLPPEYITPPSPFRVSSGSGISVSSVQDRPEEPV

```

LEDEOQLEKLPVTFEDKRENERGNLELEKRRQALLLEOKREBLAQLEBAEGR
KREBEROEKREKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEK
ENRQOELNQRKEDIVLAKAKKLEFELBALDKKQLEKODIRLUTQRO
EISTNKRELRLEITHLOOLOESQOMGRLLPEKQILNDQKQVQNSLHDSLT
TLKRALKEKELARQHLRQDLBEVETRSKLOEIDIFENQKLEIRELHNKQLOKOS
MEARLKKEKEREKITELEKOKERQORQORQOMELHQBDEHQRPLHHEEK
KREESVKRDEKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ
KVAVYRALYPFESRSHDEITIQPGDIWVKEMWDESGOTGPGMLGGELKCTGMPA
NYAKIPIENEVAPVYPKPTDSTSAAPKLARETPALVATSSSPSTPPNNMAFSS
WPTNKEPRTDMDNMAAOPSLTVPSAGQIRQSPATTPATATSSSPSPNNMAFSS
GLQOALYPRKAKDNHNFKNQVITVLEQODMWFGVQGGQKPFKSVYKLIISGP
IRKSTSMDSGSSPASKRVASPAKRVASGEFIAMVYESSGQDLTFQOQDVIL
VTKXKDMWTGTVDKAGVPSNYVRLDSESGTAGTSGISLKKPELAIQVIASTYAT
GPEQLTLAPGQILIRKKNPGGMMWEGELQARKKROIGWPPANVYKLIINRQTSITPT
EPKSTALAAVCOVIGMYDYTAQNDDELAFNKGQIINVINKEDPDWKGVEVQGLF
PSNYVKLTITMDPSQO"

misc_feature
/note="encodes EH domain"
167..406
misc_feature
/note="encodes EH domain"
767..1936
misc_feature
/note="encodes EH domain"
2324..2524
misc_feature
/note="encodes SH3 domain"
2843..3019
misc_feature
/note="encodes SH3 domain"
3110..3286
misc_feature
/note="encodes SH3 domain"
3326..3520
misc_feature
/note="encodes SH3 domain"
3569..3748
/note="encodes SH3 domain"

ORIGIN
Alignment Scores:
Pred. No.: 2,67e-31 Length: 5287
Score: 320.50 Matches: 61
Percent Similarity: 92.42% Conservative: 0
Best Local Similarity: 92.42% Mismatches: 5
Query Match: 96.25% Indels: 0
DB: 9 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x AF064243 (1-5287)
QY 1 ValysValValTyTyTyrAlaLeuTyTyrProPheGluSerArgSerHisAspGluIle 20
DB 2324 GTAAAAAGTGATATACCGGGCACTGTCCTTTGATCCGAAAGCCATGATGAATC 2383
QY 21 ThrIleGlnProGlyAspIleValMet-----ValAspGluSerGlnThr 35
DB 2384 ACTATCCAGCCAGAGACATAGTCATGTAAAGGGGAATGGGTGATGAAGCCAAACT 2443
QY 36 GlyIuProGlyTyrPheGlyGlyGlyLeuLeuTyTyrPheThrGlyTyrPheProAlaAsn 55
DB 2444 GGAGAAACCCGGCTGCTGGAGAGAAATTAAGAAAGCAAGGGTGTCCCTGCAAC 2503
QY 56 TyrAlaGluIlePro 61
DB 2504 TATGCAGAGAAATCCCA 2521

RESULT 11
AF114488 5381 bp mRNA linear PRI 16-JUL-2002
DEFINITION Homo sapiens intersectin short isoform (ITSN), complete cds.
ACCESSION AF114488
VERSION AF114488.1 GI:4808824
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Euteleostomi; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 5381)
Pucharcos, C., Fuentes, J.J., Casas, C., de la Luna, S., Alcantara, S.,
Arbones, M.L., Soriano, E., Estivill, X. and Pritchard, M.
Alu-splice cloning of human Intersectin (ITSN), a putative

```

```

multivalent binding protein expressed in proliferating and
differentiating neurons and overexpressed in Down syndrome
Eur. J. Hum. Genet. 7 (6), 704-712 (1999)
MEDLINE 99415290
JOURNAL PUBMED 10482960
REFERENCE
2 (bases 1 to 5381)
Pucharcos, C., Fuentes, J.J., Pritchard, M. and Estivill, X.
AUTHORS
Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
TITLE
Cancer Research Institute, I/Hospital de Llo., Avila.
JOURNAL
Castelldefels km. 2,7, Barcelona 08907, Spain.
FEATURES
Source
Location/Qualifiers
1..5381
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chiromosome="21"
/map="21q22.1-q22.2"
1..5381
/gene="ITSN"
269..3931
/gene="ITSN"
/codon_start=1
/product="Intersectin short isoform"
/db_xref="GI:4808825"
/translation="MAQFPFPGSGSLDIATVFERPAKHQOQHSILKPISGITTDGA
RNFQSGLPQVLAQIMALAMNNGRNDQVFSIAMKILIKLQGYQPSALPVY
KOOPVATSSAPAFMGSLAMPPLVAVAPMGSLIVGKSPVLSVPAAPVPLN
GAPVYQPLPAFAHPAATLPKSSFSRSGSQNLKLOKASFDVASVPAVMAVP
QSRILKRLQPLFNSHDKTMSGHLGPARLITLMQSLPOQALAIWISDIDQKTLA
EERILAMHLIDVMSGQPLPVLPPEYIPSPFRVRSQSGISVTSVSPVRLPEEY
LEDEOQLEKLPVTFEDKRENERGNLELEKRRQALLLEOKREBLAQLEBAEGR
KREBEROEKREKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEKOLEK
ENRQOELNQRKEDIVLAKAKKLEFELBALDKKQLEKODIRLUTQRO
EISTNKRELRLEITHLOOLOESQOMGRLLPEKQILNDQKQVQNSLHDSLT
TLKRALKEKELARQHLRQDLBEVETRSKLOEIDIFENQKLEIRELHNKQLOKOS
MEARLKKEKEREKITELEKOKERQORQORQOMELHQBDEHQRPLHHEEK
KREESVKRDEKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ
KVAVYRALYPFESRSHDEITIQPGDIWVKEMWDESGOTGPGMLGGELKCTGMPA
NYAKIPIENEVAPVYPKPTDSTSAAPKLARETPALVATSSSPSTPPNNMAFSS
WPTNKEPRTDMDNMAAOPSLTVPSAGQIRQSPATTPATATSSSPSPNNMAFSS
GLQOALYPRKAKDNHNFKNQVITVLEQODMWFGVQGGQKPFKSVYKLIISGP
IRKSTSMDSGSSPASKRVASPAKRVASGEFIAMVYESSGQDLTFQOQDVIL
VTKXKDMWTGTVDKAGVPSNYVRLDSESGTAGTSGISLKKPELAIQVIASTYAT
GPEQLTLAPGQILIRKKNPGGMMWEGELQARKKROIGWPPANVYKLIINRQTSITPT
EPKSTALAAVCOVIGMYDYTAQNDDELAFNKGQIINVINKEDPDWKGVEVQGLF
PSNYVKLTITMDPSQO"

ORIGIN
Alignment Scores:
Pred. No.: 2,72e-31 Length: 5381
Score: 320.50 Matches: 61
Percent Similarity: 92.42% Conservative: 0
Best Local Similarity: 92.42% Mismatches: 0
Query Match: 96.25% Indels: 5
DB: 9 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x AF114488 (1-5381)
QY 1 ValysValValTyTyTyrAlaLeuTyTyrProPheGluSerArgSerHisAspGluIle 20
DB 2486 GTAAAAAGTGATATACCGGGCACTGATCCCTTTGATCCGAAAGCCATGATGAATC 2545
QY 21 ThrIleGlnProGlyAspIleValMet-----ValAspGluSerGlnThr 35
DB 2546 ACTATCCAGCCAGAGACATAGTCATGTAAAGGGGAATGGGTGATGAAGCCAAACT 2605
QY 36 GlyIuProGlyTyrPheGlyGlyGlyLeuLeuTyTyrPheThrGlyTyrPheProAlaAsn 55
DB 2606 GGAGAAACCCGGCTGCTGGAGAGAAATTAAGAAAGCAAGGGTGTCCCTGCAAC 2665
QY 56 TyrAlaGluIlePro 61

```

Db 2666 TATGCGAGAGAAATCCCA 2683
RESULT 12 4458 bp DNA linear PAT 17-JUL-2003
BD205034
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205034.1 GI:33014804
VERSION BD205034.1
KEYWORDS JP 2002511267-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 2 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL
COMMENT OS Homo sapiens (human)
PN JP 2002511267-A/2
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R KORENBERG, XIAO NING CHEN
PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
PC C1201/68,
PC G01N33/68//A61K48/00 C12N15/00, C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
CC and utilization thereof.
FH Location/Qualifiers
FT source 1..5458
FT /organism="Homo sapiens (human)".
FT /db_xref="taxon:9606"
FEATURES
source 1..5458
Location/Qualifiers
1..5458
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. NO.: 2.76e-31 Length: 5458
Score: 320.50 Matches: 61
Percent Similarity: 92.42% Conservative: 0
Best Local Similarity: 92.42% Mismatches: 0
Query Match: 96.25% Indels: 5
DB: 6 Gaps: 1
US-09-720-934-2_COPY_740_800 (1-61) x BD205034 (1-5458)
QY 1 ValysvalValTtYtTtYtArgAlaLeuYrPrpPheGluSerArgSerHisaspGluLe 20
Db 2484 GTAAAGGGGTATTATACCGCGACCTGTAACCCCTTGATTCAGAAAGCCATATGAAAC 2543
QY 21 ThrilegInProGjIYAspIleValMet-----ValaspGluSerGlnThr 35
Db 2544 ACTATCCAGCCCGAGACATATCATGCTTAAAGGGGAATGGGTGGATGAAAGCCAAACT 2603
QY 36 GijyIuProGjIYTriPleuGjIyIyGluLeuLyGjIyTThGjIYTriPheProAlaasn 55
Db 2604 GGAGAACCCGGCTGGCTTGAAGAGAAATTAAGAAAGAACAGAGGGTGTCCCTGCAAC 2663
QY 56 TyraIaGluLyLePro 61
Db 2664 TATGCGAGAGAAATCCCA 2681
RESULT 13
AF114487
LOCUS AF114487 6439 bp mRNA linear PRI 16-JUL-2002

DEFINITION Homo sapiens intersecin long isoform (ITSN) mRNA, complete cds.
ACCESSION AF114487
VERSION AF114487.1 GI:4808822
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Pucharcos, C., Fuentes, J.J., Casas, C., de la Luna, S., Alcantara, S.,
Arbones, M.L., Soriano, E., Estivill, X. and Pritchard, M.
TITLE Multi-valent binding of human Intersectin (ITSN), a putative
differentiating neurons and overexpressed in Down syndrome
Eur. J. Hum. Genet. 7 (6), 704-712 (1999)
JOURNAL MEDLINE
99415290
PUBMED 10482960
REFERENCE
AUTHORS Pucharcos, C., Fuentes, J.J., Pritchard, M. and Estivill, X.
TITLE Direct Submmission
JOURNAL Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
Cancer Research Institute, L'Hospitalet de l'lo., Avia.
Castelldefels km. 2,7, Barcelona 08907, Spain
FEATURES
source 1..6439
Location/Qualifiers
1..6439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1..6439
/gene="ITSN"
269..5434
/gene="ITSN"
/codon_start=1
/product="Intersectin long isoform"
/protein_id="A029952.1"
/translation="MAQPTPFGSLDITVAITVERAKHQDSFLKPGISFTITGDQA
RNFPQSGLDPPVLAQIWMALNNDNDMDQVDFSIIMKLTKLQGVLSALPVM
KQPVATSSAPAPGMSGIIASMPPLTAVAPVGMSPITVVSVPITAAVPLAN
GAPVITDLPAPFAPRANTLPKSSPSRSGPSGLNTKQKOSDGVASVPAEAVP
QSSRLKRYQLFNHSDKWTSGHITGPQARTIMOSLSPOAQLASIMNLSIDDDGLTA
BEFTIAMHLIDVAMSGPLPVLPEVTPSFRVRSAGISVTSSTVDORLPREPV
LEDEQQLLEKLPVTYFEDKRENFERNGLLEKORQALLERQERQSLAQURERQZ
RKERROBERKQLEKLEKQERLEKQERLEKQERLEKQERLEKQERLEKQERLE
RNRRQELNQRNEQEDIVVAKAKKTLELEALNDKQQLDQKQOQNSIHRDQK
EISTNKREIRIETITLQOOLQSQMLQKQLEIDIENNOLKEIRHNKQOLQKQK
TLKRLAKELARQHLRQDLBEVETSKQLEIDIPNNOLKEIRHNKQOLQKQK
MEAEVKQKQERKQLEKQERLEKQERLEKQERLEKQERLEKQERLEKQERLE
KREESVKKQDEKQERQADQKGLRFLPHQBPAPVAPVAPVAPVAPVAPVAPVAPV
KVYVRYALYPESRSHDEITDQDPIVWKGWDESGTGEQWLGELKQKGTGEPFA
NYAEKIPENEVAPVPTDSTSAAPKALRETPAPLAVTSSESTPNNMADSSST
WPTSTNEPDEPNTNMAAPSLITVPSAGQFORSAFTPATVSGSPSVLQKQKVE
GIQDALPWPRAKQDNHLPNKNDIVITVLEQDMMWFGVQKQKMPKSYVLLSGP
IKRSTMSQSGSESPASLAKVAPAKVYSGEETIAMTYTSSSGQDLTPQGVVIL
VTXKQDMWTGTGVGKAGVPSNNVRLKQDSGSGTAGCTGKQRPETIAQVASTAT
GPEOLTLAPGOLILIRKXNPGMWGELQARKKRQIQWPEANNVYKLSPGSKITPT
BPKKSTALAAVOVIGMTDYTAQNDDELAFNKGQILINLNKEDPMMWKEVNGVGLF
AOLPMPQPIRPSQNLNGALILQCKTBAEPFKFVRYRLANDPCKQKMPKSSFLIKP
FOKPLMESBLTKEKVMATFVWKEKLTMCNTILKALRYRKMGSGKMPKVGKIGILS
MORVRYPLIKNILENTPENHPDHSKHALERKABELSCVNEGYREKRENSDPLEMI
QAHVQCEGSEQLVNSVTNCGPKRFLHSGLYAKSKKELYGFLNDFLLQTLTK
PLGSGGTQVSPKSNLYQKRYKTPILFNEVYVVKLPDPSGDEPFLFHSIDRVTYLR
ASINERTPMDVQKTAASELYETEKKKREKRYLVASQAPATIGRLMNVAVGIEAKP
CRSHGKNPQCVTMGSOCHITKTITQDILNPKWNSGCPFTIDLBQEVLCITVFEEDQ
FSPDPLGRTEIRVADIKDQSKKPVYTKCLLHVEYPTGEIYVRDLQLFDEP"

[illegible]

```

misc_feature      /note="encodes EH domain"
                  2324. .2524
misc_feature      /note="encodes SH3 domain"
                  2843. .3019
misc_feature      /note="encodes SH3 domain"
                  3110. .3286
misc_feature      /note="encodes SH3 domain"
                  3326. .3520
misc_feature      /note="encodes SH3 domain"
                  3569. .3748
misc_feature      /note="encodes SH3 domain"
                  3836. .4390
misc_feature      /note="encodes GDS domain"
                  4649. .4819
misc_feature      /note="encodes PH domain"
                  4895. .5143
misc_feature      /note="encodes C2 domain"

ALIGNMENT SCORES:
red. No.:          3.74e-31      Length:      7247
Score:             320.50        Matches:     61
Percent Similarity: 92.42%       Conservative: 0
Best Local Similarity: 92.42%    Mismatches: 0
Query Match:       96.25%        Indels:      5
DB:                9            Gaps:         1

```

QY		1	VallVysValIValTYrTyrArlgAlaleuTYrProPheGluSerArgSerHisAspGluIle	20
Dd		2324	GTRAAAGTGSTGTATTACCGGGCACTGRHCCCCCTTTGAATCCAGAAGGCATGATAAATC	2383
QY		21	THRTLeGlnProGlyAspIleValMet-----ValAspGluSerGlnThr	35
Dd		2384	ACTATCCACCAGAGACACTAGTCATGTGTTAAAGGGGAATGGGTGATGGAAAGCCAACCT	2443
QY		36	GlyGluProGlyTyrTrpLeuGlyGlyLysLeuLeuLysGlyLysTyrGlyTyrPheProAlaAsn	55
Dd		2444	GGAGAACCCGGCTGGCTTGAGAGAAATTAAAAGGAAGACAGGGTGGTTCCCTGCCAAAC	2503
QY		56	TyrAlaGlnLysIlePro	61
Dd		2504	TATGCAGAGAAATCCCA	2521
RESULT 15				
LOCUS		AF032118		
DEFINITION		Xenopus laevis intersectin mRNA, complete cds.	4103 bp	VR T 13-NOV-1998
ACCESSION		AF032118		
VERSION		AF032118.1	GI:2642624	

